ABSTRACT

DONOHUE, KEVIN VINCENT. Genomics of Tick Reproduction and Development. (Under the direction of R. Michael Roe).

The major hemelipoglyco-carrier protein (CP) found throughout the development of male and female adult American dog ticks, *Dermacentor variabilis* (Say) was sequenced. DvCP is a single transcript coding for two protein subunits that together contain three motifs—(a) a lipoprotein n-terminal domain that is a common attribute of proteins that bind lipids, carbohydrates and metals, (b) a domain of unknown function characteristic of proteins with several large open beta sheets and (c) a von Willebrand factor type D domain near the carboxy-terminus apparently important for multimerization. These motifs also found in tick vitellogenin are not shared by heme-binding proteins studied thus far in other hematophagous insects. DvCP message was highest in fat body and salivary gland but was also found in midgut and ovary. Expression was initiated by blood feeding in virgin females and not by mating typical of tick vitellogenin (Vg); and the message was found in fed males at levels similar to part fed, virgin females. CP appears to be highly conserved among the Ixodida and shares a common origin with Vg. In the second part of this study, the tick synganglion transcriptome was studied by pyrosequencing to identify neuropeptides that regulate reproduction and development. Here we characterize fourteen putative neuropeptides (allatostatin, insulin-like peptide, ion-transport peptide, sulfakinin, bursicon alpha/beta, eclosion hormone, glycoprotein hormone alpha/beta, corazonin, four orcokinins) and five neuropeptide receptors (gonadotropin receptor, leucokinin-like receptor, sulfakinin receptor, calcitonin receptor, pyrokinin receptor) from the synganglion of female American dog ticks. Many of these neuropeptides have not been previously described in the Chelicerata. An

insulin receptor substrate protein was also found indicating that at insulin signaling network is present in ticks. A putative type-2 proprotein processing convertase was also sequenced that may be involved in cleavage at monobasic and dibasic endoproteolytic cleavage sites in prohormone peptides. Quantitative real-time PCR was used to monitor developmental expression of these genes during adult female reproduction. Their physiological role during adult tick blood feeding and reproduction is be discussed.

Genomics of Tick Reproduction and Development

by
Kevin Vincent Donohue

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Entomology

Raleigh, North Carolina

2008

APPROVED BY:

Dr. R. Michael Roe Chair of Advisory Committee	Dr. Daniel E. Sonenshine
Dr. Charles A. Apperson	Dr. Christina M. Grozinger

BIOGRAPHY

Kevin Vincent Donohue was born in Albany, New York in 1974. At the age of 10 he decided to become an entomologist. He received a Bachelors of Science in Resource Management in 1996 from the State University of New York at Syracuse, College of Environmental Science and Forestry and a Masters of Science in Entomology in 2005 from North Carolina State University.

His interests during college shifted to music and he became a professional guitarist, touring the eastern US, releasing several recordings, and teaching lessons. During this time he also pursued work in the culinary arts, training with Certified Master Chef Dale L. Miller in New York. In the Spring of 2001 his girlfriend and he traveled throughout Western and Eastern Europe for three months. Upon returning to Chapel Hill, NC he decided to return to science and completed a Masters of Science in Entomology at NCSU in the lab of Dr. R. Michael Roe and continued in the same lab to pursue a Doctor of Philosophy degree in entomology with a minor in biotechnology.

ACKNOWLEDGEMENTS

I would like to acknowledge the help and support of several individuals that played an integral role in this project, especially my committee members, Dr. R. Michael Roe, Dr. Daniel E. Sonenshine, Dr. Charles A. Apperson and Dr. Christina M. Grozinger. Several other faculty and staff at NCSU also helped to make this research possible. Dr. Nigel Deighton provided proteomics and genomics support especially with GS FLX sequencing. Megan R. Ulmer and Dr. Jenn Schaff assisted with cDNA library preparation for pyrosequencing. Dr. Mark Burke and Dr. Betsy Scholl provided bioinformatics support for 454 contig assembly and batch blasting. Dr. Chris Smith wrote customized PHP and PERL scripts used to manipulate 454 data used for microarray design. I would also like to thank Drs. Sayed M.S. Khalil, Deborah M. Thompson and Joanna A. Miller for their initial guidance with molecular biology techniques. Finally, I thank all of the dedicated teachers I have had throughout my graduate career whose lessons continue to inspire me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Heme-binding storage proteins in the Chelicerata.

Molecular characterization of the major hemelipoglycoprotein in ixodid ticks

Heme-binding storage proteins in the Chelicerata

Abstract

Lipoglycoproteins in the Chelicerata that bind and store heme represent a unique evolutionary strategy to both mitigate the toxicity of heme and utilize the molecule as a prosthetic group. Knowledge of heme-binding storage proteins in these organisms is in its infancy and much of what is known is from studies with vitellogenins (Vg) and more recently the main hemolymph storage protein in ixodid ticks characterized as a hemelipoglyco-carrier protein (CP). Data have also been reported from another arachnid, the black widow spider, *Latrodectus mirabilis*, and seem to suggest that the heme-binding capability of these large multimeric proteins is not a phenomenon found only in the Acari. CP appears to be most closely related to Vg in ticks in terms of primary structure but post-translational processing is different. Tick CP and *L. mirabilis* high density lipoprotein 1 (HDL1) are similar in that they consist of two subunits of approximate molecular weights of 90 and 100 kD, are found in the hemolymph as the dominant protein, and bind lipids, carbohydrates and cholesterol. CP binds heme which may also be the case for HDL1 since the protein was found to contain a brown pigment when analyzed by native polyacrylamide gel electrophoresis. Vgs in ticks are composed of multiple subunits and are the precursor of the yolk protein, vitellin. The phylogeny of these proteins, regulation of gene expression and putative functions of binding and storing heme throughout reproduction and development are discussed. Comparisons with non-chelicerate arthropods are made in order to highlight the evolutionarily distinct mechanisms and putative functions of heme-binding storage proteins and their possible critical function in the evolution of hematophagy.

Introduction

In the Arthropoda, heme is a widely employed iron-containing molecule that serves as a prosthetic group in hemoproteins such as cytochromes (Feyereisen 1999), glycoproteins (Duncan et al. 1999b, Donohue et al. 2008), hemoglobins (Weber et al. 1985, Trotman et al. 1993, Burmester et al. 2006), nitrophorins (Weichsel et al. 1998), peritrophic matrix proteins (Devenport et al. 2006), vitellogenin/vitellin (Logullo et al. 2002; Thompson et al. 2007) as well as many novel heme-binding proteins (Oliveira et al. 1995). Free heme is cytotoxic and may result in the formation of reactive oxygen species that ultimately can lead to lipid peroxidation (Hamza et al. 1998) and result in the breakdown of lipids, carbohydrates, proteins and nucleic acids (Gutteridge and Smith 1988, Graça-Souza et al. 2006). In addition to being a vital functional component of many hemoproteins, the toxicity of heme is abated by proteins that can sequester heme. In blood-feeding arthropods that ingest relatively large amounts of blood, the capacity to sequester and store heme is vital. Insects and ticks have evolved multiple strategies to make use of heme while at the same time mitigating its toxicity. Comparisons with insect and crustacean systems are made to the extent needed to address the putative functional and evolutionary relationships of the heme-binding storage proteins in the Chelicerata.

Heme biosynthesis

The biosynthetic pathway of heme begins with the condensation of glycine and succinyl-CoA by the rate limiting enzyme 5-aminolevulinate synthase (ALAS) (EC 2.3.1.37).

Ruiz de Mena et al. (1999) were the first to characterize an invertebrate ALAS by studying the regulation of the gene in *Drosophila melanogaster*. In comparison, mammals contain two genes, *alas1* and *alas2*, the latter of which through alternative splicing yields two isoforms (Conboy et al. 1992). In humans, *alas2* is erythroid-specific and contains iron responsive elements (IRE) in the 5' UTR that allow intracellular control, while *alas1* is considered the housekeeping form that is ubiquitously expressed (Ruiz de Mena et al., 1999; Surinya et al. 1997). Furthermore, ALAS1 is transcriptionally and translationally repressed by heme (Yamamoto et al. 1982; Yamamoto et al. 1983; Furuyama et al. 2007) but ALAS2 is not. The *Drosophila* form is analogous to human *alas1* in that it does not contain iron responsive elements but is negatively regulated by heme. The promoter region of *alas* in *Drosophila* was found to contain two major transcription initiation sites, the -84 site used primarily during embryogenesis and the -60 site utilized mainly during the adult stage. These data suggest that the regulation of the enzyme and heme biosynthesis overall may be controlled by different mechanisms in adults and embryos.

A functional heme biosynthetic pathway is also present in the hematophagous insect, *Rhodnius prolixus*, but not in the southern cattle tick, *Rhipicephalus (Boophilus) microplus* (Braz et al. 1999). Braz and coworkers (1999) injected [^{14}C]ALA into *Rho*. *prolixus* and *Rhi*. *microplus* which resulted in radiolabeled heme in *Rho*. *prolixus* hemolymph and ovaries but no radiolabeled heme in *Rhi*. *microplus.* Furthermore, δ-aminolevulinate dehydratase (ALAD) (EC 4.2.1.24) activity was detected in the insect but not in the tick, and inhibition of ALAD with succinyl acetone prevented oviposition in the insect but had no effect on ticks.

These studies *in toto* suggest that ticks do not have the ability to synthesize heme, although the work is limited to a single tick species and does not consider soft ticks.

One explanation for the apparent lack of ability to synthesize heme in ticks is their absolute dependency as obligate ectoparasites on blood during all stages of development. They ingest large amounts of blood and are dependant entirely upon host-derived heme. In contrast, mosquitoes only blood feed as adults and therefore must be able to synthesize heme during the larval stages; further evidence for this is the presence of the *alas* gene in *Culix pipiens quinquefasciatus* and *Aedes aegypti* (Genbank accession numbers XP_001849249 and XP_001658653, respectively). Nevertheless, mosquitoes do recycle iron from a blood meal as will be discussed in more detail later (Zhou et al. 2007). That some groups of hematophagous arthropods can synthesize heme while others cannot highlights the variable physiological strategies used in the adaptation to blood feeding, which has apparently taken place multiple times within the Arthropoda (Law et al. 1992; Mans and Neitz 2004). Loss of the heme biosynthetic pathway has also been demonstrated in the nematode, *Coenorhabditis elegans*, as well as several other species of parasitic nematodes (Rao et al. 2005) indicating that this phenomenon is not limited to the Acari. *C. elegans* for example are able to retrieve heme from the bacteria that serve as their food source. Duncan et al. (1999a) demonstrated the presence of ALAS in the horseshoe crab, *Limulus polyphemus*, suggesting that other organisms within the Chelicerata have an intact heme biosynthetic pathway. Since *Rhi. microplus* lacks the ability to produce heme *de novo*, an alternate mechanism presumably evolved to sequester and store heme for the extended periods that ticks must survive between

bloodmeals and to transfer heme to developing oocytes. More studies are needed in ticks to determine whether the inability to synthesize heme is common within the Acari.

Molecular biology of heme-binding storage proteins in chelicerates

In male and female ticks the major hemolymph storage protein that binds heme is a hemelipoglyco-carrier protein referred to as CP (Guderra et al. 2001, 2002a, b) or HeLp (Maya-Monteiro et al. 2000, 2004). Guderra et al. (2001) found that the molecular weight of CP in *Dermacentor variabilis* (DvCP) was 200 kDa and 340 kDa by native polyacrylamide gel electrophoresis (PAGE) and gel filtration chromatography, respectively, while Maya-Monteiro et al. (2000) reported an average molecular weight of CP from *Rhi. microplus* (RmCP) as 354 kDa by native-PAGE (342 kDa) and gel filtration (365 kDa). In both organisms, CP consists of two subunits, 92 and 98 kDa for DvCP and 92 and 103 kDa for RmCP. Guddera et al. (2001) also partially described a similar protein in the soft tick, *Ornithodoros parkeri*, with an approximate molecular weight of 500 kDa although this protein did not appear to carry heme. The research on storage proteins and heme sequestration in soft ticks is minimal as compared to hard ticks.

CP in hard ticks carries cholesterol, phospholipids, monoacylglycerides, triacylglycerides and free fatty acids in addition to heme (Guddera et al. 2001; Maya-Monteiro et al. 2000). Maya-Monteiro et al. (2004) demonstrated that O_2 consumption of heme incubated with liposomes was significantly reduced with the addition of CP, suggesting that the heme moiety bound to CP had a reduced capacity to induce oxidative damage to

phospholipid bilayers. The authors also noted that cholesterol has been reported to serve in an antioxidant capacity (Smith 1991) suggesting that CP bound cholesterol may increase the antioxidant capability of the protein. Unpublished, preliminary work by Gudderra and Roe suggested that CP in *D. variabilis* could bind CO. If this is correct, it might be assumed that heme in CP could also bind oxygen.

A second heme-binding storage protein in ticks that has been more extensively studied is vitellogenin. Vitellogenin (Vg) is the large multimeric yolk protein precursor in arthropods. In ticks, Vg binds heme (Logullo et al. 2002). Logullo et al. (2002) also revealed the antioxidant capacity of Vg and further demonstrated that heme biosynthesis did not occur in the egg of *Rhi. microplus* as was previously shown for the adult of the same species (Braz et al. 1999). It was determined that the yolk protein vitellin (Vn) was able to bind up to 30 molecules of heme per molecule of Vn. During embyogenesis, the amount of heme bound by Vn increased suggesting that the heme not used by the embryo is transferred to other Vn molecules or possibly CP.

The molecular weight of Vg from *D. variabilis* determined by native PAGE and gel permeation chromatography was reported previously in the range of 320-486 kD (Sullivan et al., 1999; Gudderra et al., 2001, 2002a,b; Thompson et al., 2005). The size of Vn was in the range of 370-480 Kd as determined by Rosell and Coons (1991). Thompson et al. (2005) confirmed the identity of the putative Vg in hemolymph as the egg yolk protein precursor by tryptic digestion liquid chromatography tandem mass spectroscopy. Although there is considerable variation in the size of Vg and Vn in these reports, the molecular weight

exceeds that predicted from translation of the Vg message (Thompson et al., 2007) discussed in more detail later. If DvVg existed as a dimer in hemolymph, the predicted molecular weight would be 412 kd (2 times 206 where the latter is predicted from translation of the Vg message), which is in reasonable agreement with that previously reported from this tick (Sullivan et al., 1999; Gudderra et al., 2001, 2002a,b; Thompson et al., 2005). Vn from *D. variabilis* on SDS-PAGE migrated as seven major bands with molecular weights of 210, 172, 157, 111, 76.2, 58.7 and 50.8 KD. The 210 kD protein apparently arises from the uncleaved monomer. Although we can not confirm the exact oligomeric form of Vg and Vn in the American dog tick, comparing the known size of the protein based on the conceptual translation of the message (Thompson et al., 2007) with the protein size reported on native gels (Rosell and Coons, 1991; Sullivan et al., 1999; Gudderra et al., 2001, 2002a,b; Thompson et al., 2005) suggest the protein exists as a dimer in hemolymph and the egg. The proteomics of Vg and Vn is similar to CP in that the native proteins are found in the same molecular weight range but clearly different when resolved by SDS-PAGE. Vg in the hemolymph was shown to be the precursor to Vn in the egg while CP which is highly abundant in hemolymph is found in low levels in the egg as compared to Vn. Vg and Vn have not been studied at either the protein or molecular level in non-tick chelicerates.

Protein architecture and phylogeny of heme-binding storage proteins

CP and Vg share motifs that are characteristic of proteins in the large lipid transfer protein superfamily. More sequence data on a putative CP in ticks was recently made

available in Genbank (*Haemaphysalis longicornis* HlVg-C; Genbank accession number BAG12081). Previous to this Nene et al. (2004) published partial tentative consensus sequences (TC) homologous to CP from *Rhi*. *appendiculatus* and compared them with TCs from *A*. *variegatum*. Donohue et al. (2008) were the first to complete the sequence of DvCP (Genbank accession number ABD83654) and soon after an *A. americanum* CP was released (Genbank accession number ABK40086). Nene et al. (2004) were the first to indicate that multiple CP homologs were present in *Rhi*. *Appendiculatus*, and this finding was confirmed in another tick species, *I*. *scapularis* (Donohue et al. (2008). Donohue, Khalil, Sonenshine and Roe recently completed the sequence of a second Vg (Genbank accession number EU204907) and partially sequenced a second CP (Genbank accession number EU834070)(unpublished data).

A protein BLAST search (Altschul et al. 1997) with *H. longicornus* Vg-C returns a top match in the Genbank non-redundant database as CP and not Vg. We performed a phylogenetic analysis of CP sequences and vitellogenin sequences selected from the top 10 results of a protein BLAST search using the Neighbor-Joining method (Saitou and Nei 1987) in MEGA4 (Tamura et al. 2007) (Figure 1). This analysis indicates that *H*. *longicornis* vitellogenin (HlVg-C) is more closely related to CP than Vg in ticks. A previous *D*. *variabilis* CP sequence was also misidentified as a vitellogenin (Genbank accession number ABC94727). The high bootstrap value (94%) at the node between HlVg-C and other CPs lends support to this argument. An alignment of the n-termini of CP protein sequences further supports the hypothesis that HlVg-C is a CP and not Vg (Figure 2). All 6 proteins

contain a 15 residue secretion signal predicted by the SignalP server (Bendtsen et al. 2004). The nine residues after the secretion signal cleavage site are 100% identical in all of the sequences; overall for the data shown they share 45.4% and 98.1% identity and consensus, respectively. The addition of DvVg and DvVg2 to this alignment does not produce significant results (data not shown). However DvVg2 is expressed in a manner similar to DvVg (unpublished observation); transcription is initiated in the fat body and midgut only in the presence of 20-hydroxyecdysone and not by blood-feeding (discussed in more detail later).

A comparison of the motifs present in CPs and Vgs highlights some of the differences between these heme-binding proteins (Table 1). DvCP, AaCP and HlVg-C are of nearly identical length. Location of the lipoprotein N-terminal domain, the domain of unknown function 1943 and the von Willebrand type D domain are very similar. The RXXR convertase cleavage site is in exactly the same location, and the arginine-phenylalanine-proline-arginine quartet is also the same in all three proteins. Assuming that the predicted RXXR cleavage site in HlVg-C is correct, it would produce an N-terminal subunit sequence that is 100% identical to the first 12 amino acids that result in cleavage of DvCP and AaCP (Figure 3). Furthermore, it would only be different by a single residue if compared with DvCP2. DvVg and DvVg2 are longer in length by a minimum of 296 residues and the putative RXXR sites are in different locations. While Thompson et al. (2007) could not confirm the N-terminal sequence of the seven subunits of DvVg, two of the peptides (157 and 50.8 kD) were predicted based on the RXXR cleavage site. The cleavage site of the

other five Vg subunits could not be confirmed but it does demonstrate major structural differences between CP and Vg.

A recent phylogenetic analysis of full length CPs, vitellogenins and Vg-like proteins (Donohue et al. 2008) showed that CP grouped most closely with tick Vgs. Babin et al. (1999) previously showed that apolipophorin II/I, apolipoprotein B, vitellogenin and microsomal triglyceride transfer proteins (MTP) are derived from a common ancestor and all belong to the large lipid transfer protein (LLTP) superfamily. Smolenaars et al. (2007) compared vertebrate and invertebrate sequences from the LLTP superfamily that all contained a large lipid transfer module or motif located toward the N-terminus and reported three major groups: apoB-like LLTPs, MTPs and Vg-like LLTPs. CPs appear to be most closely related to the vitellogenins, and they group most closely with tick Vgs (Figure 1) when compared with crustacean and insect Vgs.

Limited data on soft tick CP are available. ESTs that represent incomplete putative CPs from the argasid ticks, *Argas monolakensis* and *Ornithodoros porcinus* (Genbank accession numbers EE149553 and CB722587, respectively) are 46.8% and 41.6% identical, respectively, to DvCP. A search of the putative translations of these ESTs against the Pfam database (version 22, July 2007 release) indicates the presence of a von Willebrand type D domain.

A recent review by Cunningham et al. (2007) specifically discussed hexamerins and other lipoproteins in the class Arachnida. Of particular interest are the hemolymph high-density lipoproteins (HDL1 and HDL2) that were characterized in the spider, *Latrodectus*

mirabilis (Cunningham et al. 2000). Typical lipophorins consist of subunits that are approximately 80 and 250 kD with a native molecular weight in the range of 500 kD. HDL1 in *L. mirabilis* was found by SDS-PAGE to contain two main subunits that were 90 and 103 kD, while HDL2 consisted of 440, 121 and 70 kD. Interestingly HDL1 was brown when resolved by native PAGE, indicative of the presence of heme. Furthermore HDL1 bound hydrocarbons, triacylglycerol, free fatty acids, cholesterol, diacylglycerol (trace amounts), phosphatidyl ethanolamine and phosphatidyl choline. HDL2 was found to contain a hemocyanin subunit and bound elemental copper. Hemocyanins are hexameric copper binding proteins found throughout the Arthropoda that serve to transport oxygen although HDL2 is also a lipoprotein. The subunits are typically within the range of 70 to 75 kD and although the native proteins are commonly found with six subunits, larger aggregates may form. Unfortunately, it is unknown whether the brown pigment that was associated with HDL1 was heme since spectral data for this protein is lacking. However, the similarities between *L. mirabilis* and CP in ticks are striking and may indicate the presence of another heme-binding storage protein in the Chelicerata. This would undoubtedly raise the question of why would a non-hematophagous chelicerate possess a heme-binding storage protein. Whether these spiders or other Chelicerates can synthesize heme is currently unknown. *L. mirabilis* HDL1 sequence data are not available for comparison.

The mechanism of the association of heme with CP and Vg in the Chelicerata is unknown. Thompson et al. (2007) reported that RmVg contained six repeats of P(T/P)HH(K/E)(Y/P) and DvVg contained eleven repeats of P(T/S)HH(K/E)Y at the

carboxyl terminus which are not present in any other characterized vitellogenin. These repeats were not found in DvCP suggesting that they are not critical for heme binding.

Regulation of Vg and CP gene expression

What is known about the control of transcription of heme-binding storage proteins in the Chelicerata is obtained from studies with ticks. Tick CP and Vg protein precursors are similar in structure; however their expression profiles are not (Thompson et al. 2005; Donohue et al. 2008).

To better understand the regulation of CP during adult development, Donohue et al. (2008) used a 909 bp region of the CP gene to examine its developmental regulation using Northern blots. From these studies, it was apparent that the expression of CP in whole body of adult virgin females of *D. variabilis* was initiated after attachment to the host and with the initiation of blood feeding, and appeared to increase with host-attachment duration. After attachment for 6 d, expression of CP in female ticks was greatest in the fat body and salivary gland compared to the levels found in the ovary and midgut. These results appeared to be in agreement with the CP protein levels in the same tissues as reported previously (Gudderra et al. 2002a). The appearance of both CP mRNA and protein in the salivary gland was surprising and suggests that CP might be secreted in tick saliva, a topic that is discussed in more detail later.

Vg appears to be produced from multiple tissues but not by the salivary glands as was the case for CP. The salivary glands degenerate during the time of vitellogenesis. The

regulation of Vg also appears to be different from CP in that the DvVg message appears when mating occurs in adult females and not by the initiation of blood feeding as is the case for expression and synthesis of CP. Mating results in an increase in 20-hydroxyecdysone synthesis by the epidermis and an increase in the concentration of 20-hydroxyecdysone in the hemolymph (Lomas et al. 1997; Roe et al. 2008). Injection of the same hormone into part-fed, virgin ticks attached to the host resulted in increased levels of Vg transcription and the production of vitellogenic eggs similar to that produced by replete mated females in *D. variabilis*. Apparently, even though DvCP and DvVg may share a common evolutionary origin and have similar functions in binding lipids, carbohydrates and heme, the whole body tissue sources and regulation of transcription are different.

It is interesting that blood feeding initiates transcription of CP in part-fed virgin females, but Gudderra et al. (2001) found that mating and blood feeding to repletion reduced the levels of CP protein and Thompson et al. (2005) found that mating and blood feeding to repletion increased the Vg message, Vg hemolymph protein and yolk deposition in the egg. Apparently, in the adult stage before the initiation of vitellogenesis, CP is the major tick storage protein. Then this function shifts to Vg in replete ticks as a result of mating and with increased levels of ecdysteroids. CP has been found in other tick species including *Amblyomma americanum* (Genbank accession number EF050790), *Amblyomma variegatum*, *Rhipicephalus appendiculatus*, *Rhi. microplus*, and *Ixodes scapularis* (Nene et al. 2004, Donohue et al. 2008). Other than CP and Vg, no other storage proteins in chelicerates have been specifically shown to bind heme.

It is curious that CP and Vg share a common ancestor yet the regulation of transcription is vastly different. Upstream of the Vg gene in some insects is the presence of ecdysone response elements. It would not be surprising if they were also present in the tick genome. However, little is known about the specific regulation of CP other than the discovery that an increase in the level of mRNA occurs in response to blood feeding. Iron responsive elements (IRE) have been shown to regulate translation of tick ferritin (Mulenga et al. 2004); however when CP mRNA is present, the protein is also present, so it is unlikely that an IRE regulates translation. A nucleotide BLAST against the non-redundant Genbank database with the 5' UTR of DvCP or DvVg does not return any matches to genes that contain IREs (Donohue, unpublished data), such as those found in ferritins reported by Kopáček et al. (2003). Furthermore a direct alignment of DvCP, DvCP2 and DvVg with the 5' UTR of *Ixodes ricinus* ferritin (Kopáček et al. 2003)(Genbank accession number AF068224) does not produce a significant alignment (Donohue, unpublished data).

Role of heme-binding storage proteins in sequestration, transport and excretion

Dietary heme recycling - Insects

A number of different mechanisms have evolved in insects and ticks to ensure that free heme resulting from the digestion of host hemoglobin is sequestered and excreted as hematin. This is critical to prevent oxidative stress and tissue damage. In the insect, *Rho. prolixus*, host hemoglobin is digested, free heme crosses the gut epithelium into the hemolymph, and heme is sequestered by a 15-kDa protein, the *Rhodnius* heme-binding

protein (RHBP). The heme-binding capacity of RHBP prevents lipid peroxidation and functions as an antioxidant in the insect hemolymph (Machado et al. 1998; Oliveira et al. 1995).

In *Aedes aegypti*, Zhou et al. (2007) found that ferritin is critical for host derived iron transport and storage. After a blood meal in adult females, iron levels increased in the ovaries, and host-derived iron was transferred to eggs. A higher percentage of the iron that was bound to host transferrin in the blood meal was retained as compared to heme. However, due to the comparatively larger amount of heme in the blood meal, the majority of iron transferred to eggs and retained by the female mosquito, was from heme. No hemolymph heme-transport protein has been identified in *Ae. aegypti*. However, Devenport et al. (2006) described a peritrophic matrix protein, *Aedes aegypti* intestinal mucin 1 (AeIMUC1) that efficiently bound heme *in vivo* in the molar ratio of 12:1 (heme to AeIMUC1). The protein was undetectable until 1 h after the blood meal and reached maximal levels 4 h after feeding, despite constant levels of the AeIMUC1 mRNA during digestion. Translation was regulated by heme due to heme regulatory motifs present in the chitin binding domains of the message. These data seem to suggest a mechanism for blood meal digestion in mosquitoes that relies on a peritrophic matrix for heme sequestration after digestion. Enzymatic breakdown by heme oxygenase (Pereira et al. 2007) and release of iron from host transferrin then allows for the trafficking of iron across the gut epithelium to ferritin that serves to distribute iron via the hemolymph to other tissues in the mosquito including the ovary and eggs. Ferritin from the sand fly, *Lutzomyia longipalpis*, was recently

identified in cDNA libraries constructed from midgut (Jochim et al. 2008), and peritrophin-like proteins were found in the midgut of the sand fly, *Phleotomus papatasi* (Ramalho-Ortigão et al. 2008); however detailed information on the recycling of host iron in sand flies is currently lacking.

Dietary heme recycling - Ticks

Digestion of the bloodmeal in ticks begins in the midgut (Coons et al. 1986; Tarnowski and Coons, 1989) where extracellular lysis of host erythrocytes is initiated, presumably by secreted serine proteases (Miyoshi et al. 2007). Uptake of the bloodmeal components by the digest cells in the midgut is facilitated by receptor mediated endocytosis via fluid-phase endocytosis and clathrin-coated pits (Coons et al. 1986). Intracellular digestion by heterophagy, the fusion of primary lysosomes and endosomes to form secondary lysosomes, results in the breakdown of host hemoglobin. Liberated heme from hemoglobin becomes concentrated into dense residual bodies (Tarnowski and Coons, 1989), also recently characterized as hemosomes by Lara et al. (2003). These hemosomes serve as a sequestration mechanism for hematin to prevent reduction-oxidation reactions from forming harmful free radicals. During intracellular digestion, the digest cells are found along the basal lamina where some heme is transferred to the hemocoel. The exact transfer mechanism is not yet clear, but preliminary data suggest that a membrane-protein transporter is involved (Lara et al. 2005). As heme is then transferred to the hemocoel, it is sequestered by a hemelipoglyco-carrier protein, CP as described by Gudderra et al. (2001) and HeLp as

described by Maya-Monteiro et al. (2000). CP and HeLp are essentially the same protein as discussed earlier. After the transfer of heme to CP in the hemolymph, the digest cells which contain large amounts of residual bodies in the cytoplasm, are released from the basal lamina and are substituted by replacement cells from a common midgut stem cell line (Coons et al. 1986).

During vitellogenesis, heme is also transferred to Vg (Thompson et al. 2007). It is therefore assumed that the heme-binding capacity of Vg in ticks is critical to the delivery of heme to developing oocytes and the embryo, since the embryo cannot synthesize heme. Also, since heme is brown in color and when sequestered produces brown eggs (Boctor and Kamel 1976), this coloration may provide camouflage protection from predation in leaf litter, although this has not been proven.

In order to further understand digestion at the subcellular level, Lara et al. (2005) demonstrated the path of heme movement during and after digestion using both *in vitro* and *in vivo* methods. They observed that hemoglobin was associated with vesicles larger than those that captured albumin and postulated that this association resulted from cell-surface, hemoglobin-specific receptors. Palladium mesoporphyrin IX (Pd-mP) is an interesting heme analog that when bound to globin exhibits a shift in its fluorescence spectra. When Pd-mP-globin was absorbed into specialized vesicles and digestion occurred, the emission spectra shift suggested the presence of a mixture of both Pd-mP-globin and unbound Pd-mP. Upon release into the cytosol, the spectral shift to red suggested that Pd-mP was associated with protein again, and the spectra shift to green indicated the release of free Pd-mP to the

hemosome. As previously mentioned, the digest cells are associated with the basal lamina, and heme is transferred to the hemocoel where it eventually becomes bound to CP. The hemosomes appear to be the final subcellular location of sequestered heme destined for excretion. The majority of iron obtained from a blood meal in ticks and the major form of this iron in the excreta is hematin (Sonenshine 1993). It appears that use of hematin is a common strategy for heme excretion in the Arthropoda in general.

Heme-binding proteins in the saliva of hematophagous arthropods

Nitrophorins represent a group of proteins that were first described from the blood-feeding insect, *Rho. prolixus* (Ribeiro et al. 1993, Champagne et al. 1995), and have also been described from the bed bug, *Cimex lectularius* (Valenzuela et al. 1995). Nitrophorins found in *Rho. prolixus* designated NP1, 2, 3, 4 and 7 are members of the lipocalin gene family (Andersen et al. 2004), while nitrophorins from the bed bug are related to inositol phosphatases (Valenzuela and Ribeiro 1998); nitrophorins share no significant homology between these two species. One of their critical functions is to reversibly bind nitric oxide (NO) and deliver it to the bite site of the host in order to overcome vasoconstriction. As saliva is injected into the host, NO is released and crosses cell membranes leading to vascular relaxation and unrestricted blood flow. Furthermore, Ribeiro and Walker (1994) demonstrated that histamine can replace NO from the heme pocket and bind with high affinity. Removal of histamine from the site of injury during feeding is advantageous if the insect is to feed successfully and overcome the host response.

The similarity in function but not in mRNA or protein sequence of nitrophorins between triatomines and bed bugs adds to the theory that blood-feeding has evolved separately in multiple arthropod groups (Law et al. 1992) even within the same insect order. *Rho*. *prolixus* is in the family Reduviidae while *C*. *lectularius* is in the family Cimicidae, and both families are in the order Hemiptera. Whether other blood-feeding arthropods contain nitrophorins remains an unanswered question. The salivary glands of these insects have a distinctive red color that was first reported by Wigglesworth (1943) indicating the presence of heme in the salivary gland.

DvCP is synthesized in the salivary gland of ticks as well as fat body, midgut and ovary (Donohue et al. 2008). In the salivary gland of hard ticks, CP is an apoprotein (Gudderra et al. 2002a); in the hemolymph of all life stages heme is associated with CP and was reported in a 2:1 molar ratio (Maya-Monteiro et al. 2000). Conflicting evidence suggests that CP may or may not be a component of tick saliva (Madden et al. 2002; Donohue et al. 2008). Regardless of the presence or absence of CP in saliva, tick saliva is clear in appearance suggesting that it does not contain heme (Madden et al. 2002). If this is the case, then NO sequestration as discussed before for nitrophorins is probably not a function of CP in ticks. The possibility that the major hemolymph and whole body storage protein in ticks also has a critical function in saliva for host complementation would be novel based on the studies so far conducted for storage proteins in the Arthropoda. For these reasons, the issue of the presence of CP in saliva is worth considering in more detail.

Conflicting evidence from immunoblots suggests that CP may be in the saliva and cement cone of hard ticks. Shapiro et al. (1986) demonstrated by western blots, that tick-resistant guinea pig antibodies detected a 94 kD protein in salivary gland and cement cone of 3 and 5 d fed adult female *Rhi. appendiculatus*. They found that the 94 kD protein in the salivary gland decreased in concentration during attachment to the host over 8 d in both males and females. In another study, Shapiro et al. (1987) were able to differentiate 88, 90 and 94 kD salivary gland proteins from the same tick species by western blot with the same tick-resistant guinea pig antibodies. Gordan and Allen (1987) used *Dermacentor andersoni*-resistant rabbit sera to probe western blots of fed tick salivary gland homogenate and found that proteins in the range of 88 to 94 kD were only weakly detected. This is in contrast to the findings reported by Shapiro et al. (1986) for *Rhi. appendiculatus* in which the 94 kD band was one of the dominant proteins. Jaworski et al. (1990) detected a 90 kD *A. americanum* salivary gland protein in western blots probed with sera from tick-infested sheep. To compare antigenic proteins in the salivary glands of several ixodid ticks, they produced antibodies to 3-d fed *A. americanum* female salivary glands and then preabsorbed them with unfed salivary gland homogenate from the same tick species. These antibodies recognized a 90 kD protein in the salivary glands of *D. variabilis*, *I. dammini* and *A. americanum* suggesting that this protein was upregulated during feeding. These studies appear to suggest that related polypeptides of similar molecular weight to the subunits of DvCP appear to be involved in host complementation. The function of these proteins with similar molecular weight to the subunits of CP was not determined in these studies.

Reports of a ~90 kD polypeptide in tick cement versus saliva have been equivocal. As mentioned previously, Shapiro et al. (1986) indicated that a 94 kD protein was present in *Rhi*. *appendiculatus* cement of 3 and 5 d fed ticks. Jaworski et al. (1992) used antibodies raised against the 90 kD salivary gland protein to probe cement cone proteins in *D. variabilis*, *A. americanum* and *Rhi*. *sanguineus*. They found that these antibodies recognized a 70 kD protein in the cement of these three tick species and not a 90 kD protein, suggesting that either the antibody non-specifically bound to another protein or the 90 kD protein existed in a modified form in the cement cone. Furthermore, coomassie blue stained SDS-polyacrylamide gels of cement proteins from *D. variabilis* and *A. americanum* failed to show any protein of approximately 90 kD (Jaworski et al. 1992). Wang and Nuttall (1994) compared the protein profile of salivary glands with that of saliva in *Rhi*. *appendiculatus* and found a 98 kD protein present in unfed, 2, 4, 6 and 8 day fed female ticks and in saliva of 6 and 8 d fed female ticks. Based on equal protein loading of the gel (SDS-PAG), it appeared that the concentration of the 98 kD band increased with feeding. The 98 kD protein in saliva appeared to be a dominant component of the saliva sample as well as the dominant hemolymph protein; this protein may be CP. Antisera from denatured salivary gland extract of unfed and 6 d fed ticks failed to recognize this protein although tick-immune serum weakly recognized the 98 kD protein in hemolymph, 6 d saliva and 6 d salivary gland extract. Furthermore, results of immunoblots showing common bands in hemolymph, salivary glands and saliva but in different concentration led the authors to conclude that proteins expressed in

the salivary gland then undergo further modification before being secreted into hemolymph and saliva.

Madden et al. (2002) analyzed the saliva of *A. americanum* and *Amblyomma maculatum* and found that the dominant protein in each sample had an identical N-terminal sequence to that of RaCP (Maya-Monteiro et al. 2000) and nearly identical to that of DvCP (Gudderra et al. 2001). This argues for a functional role for CP in saliva and host complementation. However, Donohue et al. (2008) analyzed the cement cone and saliva of *D. variabilis* to determine whether DvCP found in hemolymph and salivary glands was also a component of saliva. Repeated careful collections of saliva and cement cone from *D. variabilis* analyzed by PAGE failed to show any band with similar mobility to that of hemolymph DvCP. To address whether DvCP might exist in a modified form in cement cone, a possibility suggested by Jaworski et al. (1992), cement cone proteins were separated by SDS-PAGE and subjected to tryptic digestion mass fingerprinting. A positive control of cement cone spiked with hemolymph which contains DvCP was used to demonstrate that CP could be detected by this method in a complex sample of proteins which would include proteins from the cement cone. Proteins ≥66 kD and proteins <66 kD from the soluble cement cone were resolved and subjected to tryptic digestion LC tandem mass spectroscopy. DvCP fragments could not be identified in either molecular weight range of the cement cone while DvCP could be positively identified in the positive control. These data suggest that either DvCP is not a component of tick saliva or cement cone, or it is present below the detection limit of silver stained polyacrylamide gels and that of LC-MS/MS (the latter lower

limit of detection was approximately 50-75 ng for a protein with a molecular weight of ~100 kD).

The discrepancies on whether CP is present in saliva in part may be due to the unnatural method of saliva collection, which usually involves the injection of dopamine and pilocarpine or theophylline. In some procedures, researchers injected relatively large volumes of material in multiple penetrations of the body wall into the tick hemocoel to deliver the drug. This treatment may provide an inaccurate representation of what is contained in saliva from ticks normally feeding on a host. Also, since CP is the predominant protein in tick hemolymph and saliva is usually collected externally from the mouthparts, then hemolymph contamination of the saliva is at least theoretically possible. In the preliminary work of Donohue et al. (2008), it was noted that hemolymph released from a single (3-5 microliter) dopamine/pilocarpine injection by movement over the outside surface of the cuticle was able to contaminate the tick mouth parts; in this case, CP was actually resolved from the saliva collected. Special precautions were needed to prevent this contamination. Whether this was an issue in other research efforts is difficult to determine. Donohue, Sonenshine and Roe (unpublished data) found that salivary glands held in organ culture, released CP protein into the incubation medium as determined by co-electrophoresis in separate lanes with hemolymph CP. However, it could not be determined in these studies whether the release was from the salivary ducts or directly from the salivary gland cells. Considering the large amount of salivary gland tissue at the time of blood feeding and that CP gene expression in the salivary gland is initiated at this time presumably to store digestion

products obtained from host blood, an argument can be made that the salivary glands would be a reasonable source for hemolymph CP. Also, multiple sources for the production of other storage proteins like Vg have been shown before in ticks (Thompson et al., 2005, 2007); in this case the midgut and fat body produce Vg and not the salivary glands which are degenerating during the time of vitellogenesis. On the other hand, the idea that the salivary gland contributes to the production of the predominant hemolymph proteins would be novel.

In summary, research on the presence of CP in saliva is conflicting, although it appears that most of the studies suggest that the protein is involved in some aspect of tick-host interaction. However, Donohue et al. (2008) was unable to find CP in the saliva or cement cone of the American dog tick. The reason for this absence of CP in saliva is not understood. Research is needed in multiple tick species and laboratories to resolve this issue and determine if the major hemolymph and whole body tick storage protein is also secreted into saliva for host complementation and/or is secreted from the salivary glands into the hemolymph. It is interesting to note that Gudderra et al. (2001) reported that CP also might occur in tick coxal fluid.

Conclusions and Future Directions

Vg and CP in ticks are the major hemolymph and whole body storage proteins that also sequester heme; since heme biosynthesis does not occur in ticks and ticks go through extended periods where they do not feed, it appears that heme storage may be critical for nymphal, larval, adult and embryo development and that Vg and CP may be essential to tick

hematophagy. The use of Vg and CP for this function appears to be unique within the Arthropoda but more studies are needed in the Chelicerata to understand the evolution of heme sequestration by these proteins. CP may have other functions in coloration, host complementation, excretion, and respiration. The presence of CP in saliva and coxal fluid is equivocal, although in the case of the former, most studies indicate the protein is found in saliva. Most of our knowledge of heme storage proteins and heme biosynthesis in ticks is limited to hard ticks and to a very small number of tick species. In the case of heme synthesis, the work is limited to a single species. Our knowledge of the differential regulation of CP versus Vg is minimal, with most of the work conducted on the latter, where ecdysteroids have been shown to regulate the Vg message. We also do not understand why these proteins are produced in some cases by different tissues, how heme binds to CP and Vg, or why multiple proteins of both CP and Vg occur in ticks. It appears that the major hemelipoglycoproteins of tick hemolymph and whole body represent a unique adaption to the tick life strategy of hematophagy which separates them from other storage proteins so far studied within the Arthropoda.

References

Altschul, S.F., Madden, T.L., Schäffer, 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 Research 25, 3389-3402.

Andersen, J.F., Gudderra, N.P., Francischetti, I.M.B., Valenzuela, J.G., Ribeiro, J.M.C., 2004. Recognition of anionic phospholipid membranes by an antihemostatic protein from a blood-feeding insect. Biochemistry 43, 6987-6994.

Babin, P.J., Boger, J., Kooiman, F.P., Van Marrwijk, W.J.A., Van der Horst, D.J., 1999. Apolipophorin II/I, apolipoprotein B, vitellogenin, and microsomal triglyceride transfer protein genes are derived from a common ancestor. Journal of Molecular Evolution 49, 150-160.

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

Braz, G.R.C., Coelho, H.S.L. Masuda, H., Oliveira, P.L., 1999. A missing metabolic pathway in the cattle tick *Boophilus microplus*. Current Biology 9, 703-706.

Boctor, F.N., Kamel, M.Y., 1976. Purification and characterization of two lipovitellins from eggs of the tick, *Dermacentor andersoni*. Insect Biochemistry 6, 233-240.

Burmester, T., Storf, J., Hasenjäger, A., Klawitter, S., Hankeln T., 2006. The hemoglobin genes of *Drosophila*. FEBS Journal 273, 468-480.

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

Conboy, J.G., Cox, T.C., Bottomley, S.S., Bawden, M.J., May, B.K., 1992. Human erythroid 5-aminolevulinate synthase. Journal of Biological Chemistry 267, 18753-18758.

Coons, L.B., Rosell-Davis, R., Tarnowski, B.I., 1986. Bloodmeal digestion in ticks. *In* Morphology, Physiology and Behavioral Biology of Ticks (Sauer, J.R. and Hair, J.A., eds), pp. 248-279. Ellis Horwood, Chinchester, England.

Cunningham, M., González, A., Poller, R. 2000. Characterization of lipoproteins isolated from the hemolymph of the spider *Latrodectus mirabilis* (Araneae, Theridiidae). The Journal of Arachnology 28, 49-55.

Cunningham, M., Garcia, F., Pollero, R.J., 2007. Arachnid lipoproteins: comparative aspects. Comparative Biochemistry and Physiology, Part C 146, 79-87.

Devenport, M., Alvarenga, P.H., Shao, L., Fujioka, H., Bianconi, M.L., Oliveira, P.L, Jacobs-Lorena, M., 2006. Identification of the *Aedes aegypti* peritrophic matrix protein AeIMUCI as a heme-binding protein 45, 9540-9549.

Donohue, K.V., Khalil, S.M.S., Mitchell, R.D., Sonenshine, D.E., Roe, R.M., 2008. Molecular characterization of the major hemelipoglycoprotein in ixodid ticks. Insect Molecular Biology 17, 197-208.

Duncan, R., Faggart, M.A., Roger, A.J., Cornell, N.W., 1999a. Phylogenetic analysis of the 5-aminolevulinate synthase gene. Molecular Biology and Evolution 16, 383-396.

Duncan, T., Osawa, Y., Kutty, R.K., Kutty, G., Wiggert, B., 1999b. Heme-binding by *Drosophila* retinoid- and fatty acid-binding glycoprotein (HRABG), a member of the proapolipophorin gene family. Journal of Lipid Research 40, 1222-1228.

Feyereisen, R., 1999. Insect P450 enzymes. Annual Review of Entomology 44, 507-533.

Furuyama, K., Kaneko, K., Vargas V., P.D., 2007. Heme as a magnificent molecule with multiple missions: Heme determines its own fate and governs cellular homeostasis. Tohoku Journal of Experimental Medicine 213, 1-16.

Gordan, J.R., Allen, J.R., 1987. Isolation and characterization of salivary antigens from the female tick, *Dermacentor andersoni*. Parasite Immunology 9, 337-352.

Graça-Souza, A.V., Maya-Monteiro, C., Paiva-Silva, G.O., Braz, G.R.C., Paes, M.C., Sorgine, M.H.F., Oliveira, M.F., Oliveira, P.L., 2006. Adaptations against heme toxicity in blood-feeding arthropods. Insect Biochemistry and Molecular Biology 36, 322-335.

Gudderra, N.P., Neese, P.A., Sonenshine, D.E., Apperson, C.S., Roe, R.M., 2001. Developmental profile, isolation, and biochemical characterization of a novel lipoglycoheme-carrier protein from the American dog tick, *Dermacentor variabilis* (Acari: Ixodidae) and observations on a similar protein in the soft tick, *Ornithodoros parkeri* (Acari: Argasidae). Insect Biochemistry and Molecular Biology 31, 299-311.

Gudderra, N.P., Sonenshine, D.E., Apperson, C.S., Roe, R.M., 2002a. Tissue distribution and characterization of predominant hemolymph carrier proteins from *Dermacentor variabilis* and *Ornithodoros parkeri*. Journal of Insect Physiology 48, 161-170.

Gudderra, N.P., Sonenshine, D.E., Apperson, C.S., Roe, R.M., 2002b. Hemolymph proteins in ticks. Journal of Insect Physiology 48, 269-278.

Gutteridge, J.M.C., Smith, A., 1988. Antioxidant protection by haemopexin of haem-stimulated lipid peroxidation. Journal of Biochemistry 256, 861-865.

Hamza, I., Chauhan, S., Hassett, R., O'Brian, M.R., 1998. The bacterial IRR protein is required for coordination of heme biosynthesis with iron availability. Journal of Biological Chemistry 273, 21669-21674.

Jaworski, D.C., Muller, M.T., Simmen, F.A., Needham, G.R. 1990. *Amblyomma americanum*: identification of tick salivary gland antigens from unfed and early feeding females with comparisons to *Ixodes dammini* and *Dermacentor variabilis*. Experimental Parasitology 70, 217-226.

Jaworski, D.C., Rosell, R., Coons, L.B., Needham, G.R., 1992. Tick (Acari: Ixodidae) attachment cement and salivary gland cells contain similar immunoreactive polypeptides. Journal of Medical Entomology 29, 305-309.

Jochim, R.C., Teixeira, C.R., Laughinghouse, A., Mu, J., Oliveira, F., Gomes, R.B., Elnaiem, D.-E., Valenzuela, J.G., 2008. The midgut transcriptome of *Lutzomyia longipalpis*: comparative analysis of cDNA libraries from sugar-fed, blood-fed, post-digested and *Leishmania infantum chagasi*-infected sand flies. BMC Genomics 9, 15.

Kopáček, P., Ždychová, J., Yoshiga, T., Weise, C., Rudenko, N., Law, J.H., 2003. Molecular cloning, expression and isolation of ferritins from two tick species—*Ornithodoros moubata* and *Ixodes ricinus*. Insect Biochemistry and Molecular Biology 33, 103-113.

Lara, F.A., Lins, U., Paiva-Silva, G., Almeida, I.C., Braga, C.M., Miguens, F.C., Oliveira, P.L., Dansa-Petretski, M., 2003. A new intracellular pathway of haem detoxification in the midgut of the cattle tick *Boophilus microplus*: aggregation inside a specialized organelle, the hemosome. Journal of Experimental Biology 206, 1707-1715.

Lara, F.A., Lins, U., Bechara, G.H., Oliveira, P.L., 2005. Tracing heme in a living cell: hemoglobin degradation and heme traffic in digest cells of the cattle tick *Boophilus microplus*. Journal of Experimental Biology 208, 3093-3101.

Law, J.H., Ribeiro, J.M.C., Wells, M.A., 1992. Biochemical insights derived from insect diversity. Annual Review of Biochemistry 61, 87-111.

Logullo, C., Moraes, J., Dansa-Petretski, M., Vaz Jr., I.S., Masuda, A., Sorgine, M.H.F., Braz, G.R., Masuda, H., Oliveira, P.L., 2002. Binding and storage of heme by vitellin

from the cattle tick, *Boophilus microplus*. Insect Biochemistry and Molecular Biology 32, 1805-1811.

Lomas, L.O., Turner, P.C., Rees, H.H., 1997. A novel neuropoeptide-endocrine interaction controlling ecdysteroid production in ixodid ticks. Proceedings of the Royal Society of London B 264, 589-596.

Machado, E.A., Oliveira, P.L. Moreira, M.F., de Souza, W., Masuda, H., 1998. Uptake of *Rhodnius* heme-binding protein (RHBP) by the ovary of *Rhodnius prolixus*. Archives of Insect Biochemistry and Physiology 39, 133-143.

Madden, R.D., Sauer, J.R., Dillwith, J.W., 2002. A proteomics approach to characterizing tick salivary secretions. Experimental and Applied Acarology 28, 77-87.

Mans, B.J., Neitz, A.W.H., 2004. Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. Insect Biochemistry and Molecular Biology 34, 1-17.

Maya-Monteiro, C.M., Daffre, S., Logullo, C., Lara, F.A., Alves, E.W., Capurro, M.L., Zingali, R., Almeida, I.C., Oliveira, P.L., 2000. HeLp, a heme lipoprotein from the hemolymph of the cattle tick, *Boophilus microplus*. Journal of Biological Chemistry 275, 36584036589.

Maya-Monteiro, C.M., Alves, L.R., Pinhal, N., Abdalla, D.S.P., Oliveira, P.L., 2004. HeLp, a heme-transporting lipoprotein with an antioxidant role. Insect Biochemistry and Molecular Biology 34,, 81-87.

Miyoshi, T., Tsuji, N., Islam, M.K., Huang, X., Motobu, M., Abdul Alim, M., Fujisaki, K., 2007. Molecular and reverse genetic characterization of serine proteinase-induced hemolysis in the midgut of the ixodid tick *Haemaphysalis longicornis*. Journal of Insect Physiology 53, 195-203.

Mulenga, A., Simer, J.A., Macaluso, K.R., Azad, A.F., 2004. Stress and transcriptional regulation of tick ferritin HC. Insect Molecular Biology 13, 423-433.

Nene, V., Lee, D., Kang'a, S., Skilton, R., Shah, T., de Villiers, E., Mwaura, S., Taylor, D., Quackenbush, J., Bishop, R., 2004. Genes transcribed in the salivary glands of female *Rhipicephalus appendiculatus* ticks infected with *Theileria parva*. Insect Biochemistry and Molecular Biology 34, 1117-1128.

Oliveira, P.L., Kadoka, J.K., Iberia, J.M., Meyer, T., Poorman, R., Laves, E.W., Walker, F.A., Machado, E.A., Nussenzveig, R.H., Padovan, G.J., Masuda, H., 1995. A heme-

binding protein from hemolymph and oocytes of the blood-sucking insect, *Rhodnius prolixus*. Journal of Biological Chemistry 270, 10897-10901.

Pereira, L.O.R., Oliveira, P.L., Almeida, I.C., Paiva-Silva, G.O., 2007. Biglutaminyl-biliverdin IX alpha as a heme degradation product in the dengue fever insect-vector *Aedes aegypti*. Biochemistry 46, 6822-6829.

Ramalho-Ortigão, M., Jochim, R.C., Anderson, J.M., Lawyer, P.G., Pham, V.-M., Kamhawi, S., Valenzuela, J.G., 2008. 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.

Rao, A.U., Carta, L.K., Lesuisse, E.L., Hamza, I., 2005. Lack of heme synthesis in a free-living eukaryote. Proceedings of the National Academy of Sciences USA 102, 4270-4275.

Ribeiro, J.M.C., Hazzard, J.M.H., Nussenzveig, R.H., Champagne, D.E., Walker, F.A., 1993. Reversible binding of nitric oxide by a salivary heme protein from a bloodsucking insect. Science 260, 539-541.

Ribeiro, J.M.C, Walker, F.A., 1994. High affinity histamine-binding and antihistaminic activity of the salivary nitric oxide-carrying heme protein (nitrophorins) of *Rhodnius prolixus*. Journal of Experimental Medicine 180, 2251-2257.

Roe, R.M., Donohue, K.V., Khalil, S.M., Sonenshine, D.E., 2008. Hormonal regulation of metamorphosis and reproduction in ticks. Frontiers in Bioscience 13, 7250-7268.

Ruiz de Mena, I., Fernández-Moreno, M.A., Bornstein, B., Kaguni, L.S., Garesse, R., 1999. Structure and regulated expression of the δ-aminolevulinate synthase gene from *Drosophila melanogaster*. Journal of Biological Chemistry 274, 37321-37328.

Rosell, R., Coons, L.W., 1991. Purification and partial characterization of vitellin from the eggs of the hard tick, *Dermacentor variabilis*. Insect Biochemistry 21, 871-885.

Saitou, N., Nei, M., 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4, 406-425.

Shapiro, S.Z., Voigt, W.P., Fujisaki, K., 1986. Tick antigens recognized by serum from a guinea pig resistant to infestation with the tick *Rhipicephalus appendiculatus*. Journal of Parasitology 72, 454-463.

Shapiro, S.Z., Busher, G., Dobbelaere, D.A.E., 1987. Acquired resistance to *Rhipicephalus appendicualtus* (Acari: Ixodidae): identification of an antigen eliciting resistance in rabbits. Journal of Medical Entomology 24, 147-154.

Sonenshine, D.E., 1993. Biology of Ticks, Volume 1. Oxford University Press, 447 pp.

Smith, L.L., 1991. Another cholesterol hypothesis: cholesterol as antioxidant. Free Radical Biology and Medicine 11, 47-61.

Smolenaars, M.M.W., Madsen, O., Rodenburg, K.W., Van der Horst, D.J., 2007. Molecular diversity and evolution of the large lipid transfer protein superfamily. Journal of Lipid Research 48, 489-502.

Sullivan, C.D., Rosell, R.C., Coons, L.B., 1999. Partial characterization of vitellogenin from the ixodid *Dermacentor variabilis*: preliminary results. In: Needham, G.R., Mitchell, R., Horn, D.J., Welbourn, W.C. (Eds.), Acarology IX, vol. 2, Symposia, July 17-22, 1994. Ohio Biological Survey, Columbus, Ohio, pp. 477-480.

Surinya, K.H., Cox, T.C., May, B.K., 1997. Transcriptional regulation of the human erythroid 5-aminolevulinate synthase gene. Journal of Biological Chemistry 271, 26585-26594.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596-1599.

Tarnowski, B.I., Coons, L.B., 1989. Ultrastructure of the midgut and bloodmeal digestion in the adult tick *Dermacentor variabilis*. Experimental and Applied Acarology 6, 263-289.

Thompson, D.M., Khalil, S.M.S., Jeffers, L.A., Ananthapadmanaban, U., Sonenshine, D.E., Mitchell, R.D., Osgood, C.J., Apperson, C.S., Roe, R.M., 2005. In vivo role of 20-hydroxyecdysone in the regulation of the vitellogenin mRNA and egg development in the American dog tick, *Dermacentor variabilis* (Say). Journal of Insect Physiology 51, 1105-1116.

Thompson, D.M., Khalil, S.M.S., Jeffers, L.A., Sonenshine, D.E., Mitchell, R.D., Osgood, C.S., Roe, R.M., 2007. Sequence and the developmental regulation of the first complete vitellogenin messenger RNA from ticks responsible for heme sequestration. Insect Biochemistry and Molecular Biology 37, 363-374.

Trotman, C.N.A., Manning, A.M., Bray, J.A., Jellie, A.M., Moens, L., 1993. Interdomain linkage in the polymeric hemoglobin molecule of *Artemia*. Journal of Molecular Evolution 38, 628-636.

Valenzuela, J.G., Ribeiro, J.M.C., 1998. Purification and cloning of the salivary nitrophorins from the hemipteran *Cimex lectularius*. Journal of Experimental Biology 201, 2659-2664.

Valenzuela, J.G., Walker, F.A., Ribeiro, J.M.C., 1995. A salivary nitrophorin (nitric oxide-carrying hemoprotein) in the bedbug *Cimex lectularius*. Journal of Experimental Biology 198, 1519-1526.

Wang, H., Nuttall, P.A., 1994. Comparison of the proteins in salivary glands, saliva and haemolymph of *Rhipicephalus appendiculatus* female ticks during feeding. Parasitology 109, 517-523.

Weber, R.E., Braunitzer, G., Kleinschmidt, T., 1985. Functional multiplicity and structural correlations in the hemoglobin system of larvae of *Chironomus thummi thummi* (Insect, Diptera): Hb components CTT I, CTT II, CTT III, CTT IV, CTT VI, CTT VIIB, CTT IX and CTT X. Comparative Biochemistry and Physiology 80B, 747-753.

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. Nature Structural Biology 5, 304-309.

Wigglesworth, V.B., 1943. The fate of haemoglobin in *Rhodnius prolixus* (Hemiptera) and other blood-sucking arthropods. Proceedings of the Royal Society of London Series B, Biological Sciences 131, 313-339.

Yamamoto, M., Hayashi, N., Kikuchi, G., 1982. Evidence for the transcriptional inhibition by heme of the synthesis of delta-aminolevulinate synthase in rat liver. Biochemical and Biophysical Research Communications 105, 985-990.

Yamamoto, M., Hayashi, N., Kikuchi, G., 1983. Translational inhibition by heme of the synthesis of hepatic delta-aminolevulinate synthase in a cell-free system. Biochemical and Biophysical Research Communications 115, 225-231.

Zhou, G., Kohlhepp, P., Geiser, D., Carmen Frasquillo, M., Vazquez-Moreno, L., Winserling, J.J., 2007. Fate of blood meal iron in mosquitoes. Journal of Insect Physiology 53, 1169-1178.

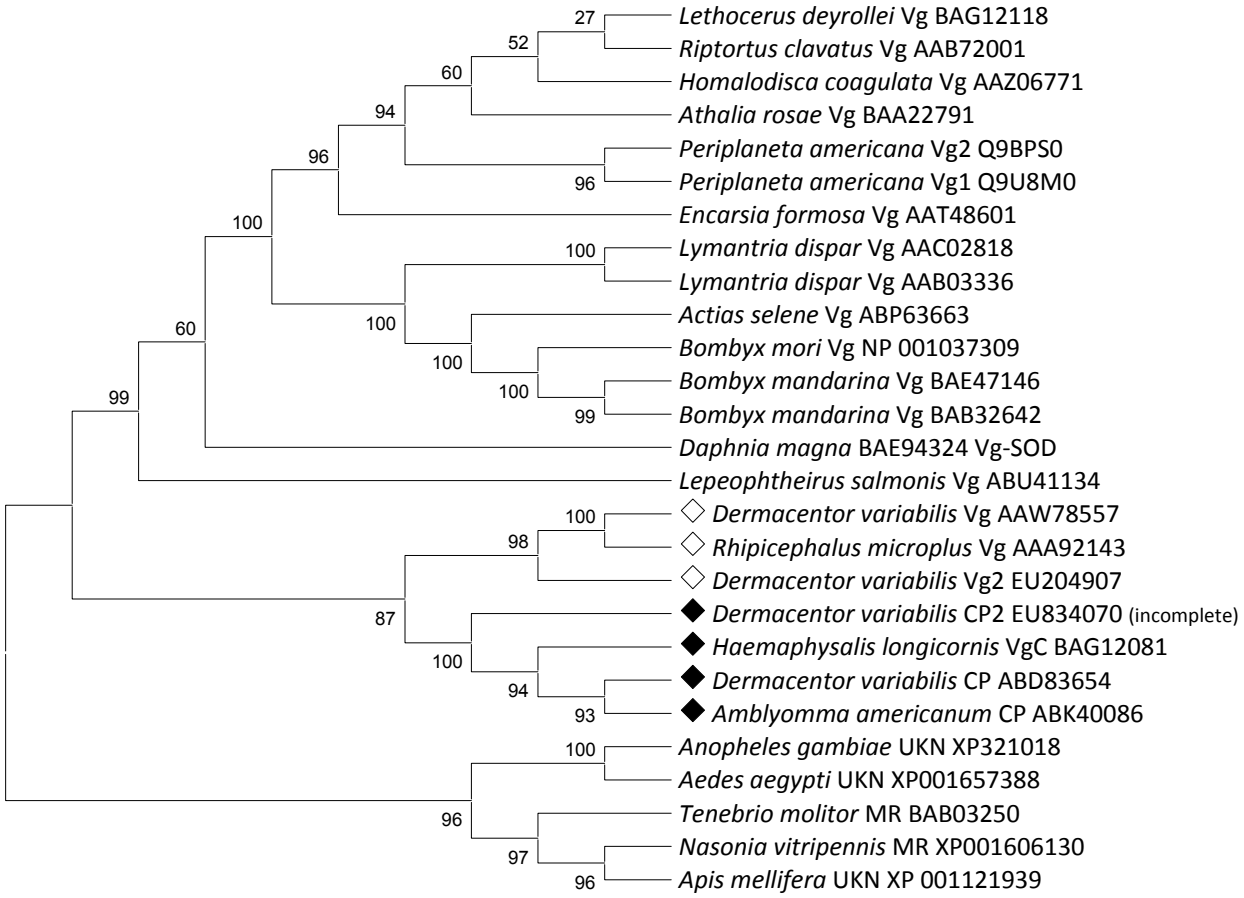


Figure 1. Evolutionary relationship of hemelipoglyco-carrier proteins (CP) and vitellogenins (Vg) with the 10 best matching amino acid sequences (to CP and to Vg separately) from the Genbank non-redundant database. BLASTp was performed with default parameters. The evolutionary history was inferred using the Neighbor-Joining method in MEGA4. The optimal tree with the sum of branch length = 14.713 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown at each node. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 856 positions in the final dataset. Tick Vgs are labeled with open diamonds, and putative and confirmed CPs are labeled with closed diamonds.

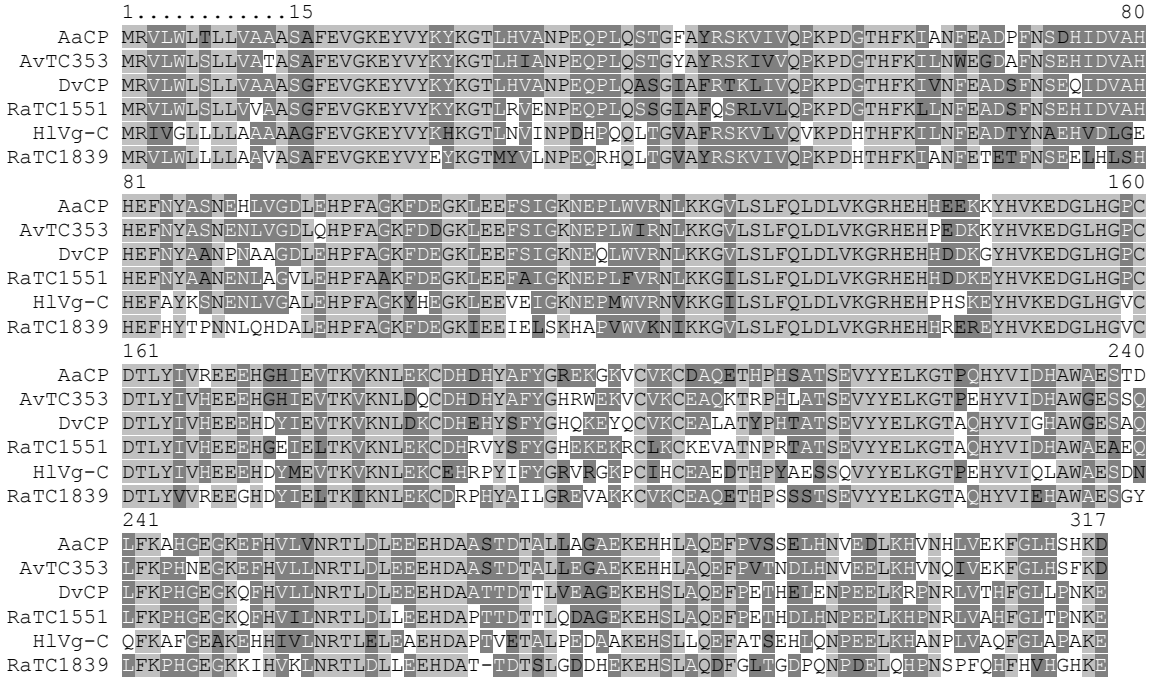


Figure 2. ClustalW alignment of the first 317 residues of hemelipoglyco-carrier protein (CP) from *Dermacentor variabilis* (DvCP), *Amblyomma americanum* (DvCP), *Amblyomma variegatum* (AvTC353), *Rhipicephalus appendiculatus* (RaTC1551 and Ra1839) and *Haemaphysalis longicornis* (HlVg-C). Residues 1 through 15 represent the predicted secretion signal.

Figure 3. ClustalW alignment of the RXXR convertase cleavage site and previously confirmed first thirty seven residues of the ~92 kD subunit of *Dermacentor variabilis* hemelipoglyco-carrier protein (DvCP) with DvCP2, *Amblyomma americanum* (AaCP) and *Haemaphysalis longicornis* vitellogenin-C (HlVg-C). The first four residues are the RXXR cleavage site.

Table 1. Comparison of the locations of known motifs in the full length amino acid sequences of tick hemelipoglyco-carrier protein (CP) and vitellogenin (Vg) from *Dermacentor variabilis* (Dv), *Amblyomma americanum* (Aa), and *Hemaphysalis longicornis* (Hl). Numbers below the lipoprotein n-terminal domain (LPD_N), domain of unknown function 1943 (DUF1943), von Willebrand type D domain (vWD_D) and RXXR convertase cleavage site indicate the predicted amino acids that conform to these motifs.

Description	GB Acc#	Length	LPD_N	DUF1943	vWD_D	RXXR site
DvCP	AAW78577	1547	16-622	667-945	1351-1522	746-749 (RFPR)
AaCP	ABK40086	1546	16-622	667-945	1360-1521	746-749 (RFPR)
HlVg-C	BAG12081	1545	16-622	658-945	1365-1524	746-749 (RFPR)
DvVg	AAW78577	1843	34-721	753-1051	1488-1656	465-468 (RLFR)
DvVg2	ABW82681	1887	28-737	769-1057	1566-1744	853-856 (RGVR)

This chapter formatted for and published in
Insect Molecular Biology 2008, 17(3):197-208.

Running Title: Molecular characterization of the major tick hemelipoglycoprotein

Send Galley Proofs to:
Dr. R. Michael Roe
Department of Entomology
Campus Box 7647
North Carolina State University
Raleigh, NC 27695-7647
919-515-4325 (phone/fax)
michael_roe@ncsu.edu

Molecular characterization of the major hemelipoglycoprotein in ixodid ticks

Kevin V. Donohue[a], Sayed M. S. Khalil[a], Robert D. Mitchell[b], Daniel E. Sonenshine[b,] and R. Michael Roe[a,*]

[a]Department of Entomology, Campus Box 7647, North Carolina State University, Raleigh,NC 27695-7647, USA
[b]Department of Biological Sciences, Old Dominion University, Norfolk, Virginia 23529, USA

Keywords: heme, lipoprotein, hematophagy, blood feeding, Ixodidae

* Corresponding author. Department of Entomology, Campus Box 7647, North Carolina State University, Raleigh, NC 27695, USA.
Tel: +1 919 515 4325; fax +1 919 515 4325
E-mail address: michael_roe@ncsu.edu (R.M. Roe)

Abstract

The major hemelipoglyco-carrier protein (CP) found throughout the development of male and female adult American dog ticks, *Dermacentor variabilis* (Say) was sequenced. DvCP is a single transcript coding for two protein subunits that together contain three motifs—(a) a lipoprotein n-terminal domain that is a common attribute of proteins that bind lipids, carbohydrates and metals, (b) a domain of unknown function characteristic of proteins with several large open beta sheets and (c) a von Willebrand factor type D domain near the carboxy-terminus apparently important for multimerization. These motifs also found in tick vitellogenin are not shared by heme-binding proteins studied thus far in other hematophagous insects. DvCP message was highest in fat body and salivary gland but was also found in midgut and ovary. Expression was initiated by blood feeding in virgin females and not by mating typical of tick vitellogenin; and the message was found in fed males at levels similar to part fed, virgin females. CP appears to be highly conserved among the Ixodida. The closest match by BLASTP to DvCP is vitellogenin from *Caenorhabditis elegans* (AAC04423) suggesting that CP is a novel protein. The role of CP in heme sequestration, the evolution of hematophagy and host complementation is discussed.

Introduction

Adaptation to hematophagy has developed multiple times in ticks and insects, even within a particular group such as the Diptera (Law *et al*., 1992; Mans *et al*., 2004). An important adaptation that co-evolved with blood feeding is heme sequesteration by heme-binding proteins and heme excretion, both of which prevent oxidative stress and tissue damage. The cytotoxicity of free heme results from the formation of reactive oxygen species that leads to lipid peroxidation (Hamza *et al*., 1998).

Heme is also important as a prosthetic group for respiration, enzymatic detoxification and oxygen transport (Furuyama *et al*. 2007). In *Rhodnius prolixus*, host hemoglobin is digested to free heme which is absorbed into the hemolymph and sequestered by a 15-kDa heme-binding protein (RHBP). RHBP prevents lipid peroxidation and functions as an antioxidant (Machado *et al*., 1998; Oliveira *et al*., 1995). Other heme-binding proteins present in *R*. *prolixus* include nitrophorins for nitric oxide transport (Weichsel *et al*., 1998) and which have been implicated in host vasodilation during blood feeding. This suggests multiple uses for heme and heme binding proteins in blood feeding insects and possibly in other organisms.

Despite the abundance of heme from the host hemoglobin, triatomines apparently have the ability to synthesize heme as evident by the expression of delta-aminolevulinic acid dehydratase, the rate limiting enzyme in the heme biosynthetic pathway (Braz *et al*., 1999). In contrast to *R*. *prolixus*, the southern cattle tick,

Rhipicephalus (Boophilus) microplus (Braz *et al*., 1999) has lost the ability to make heme. Because of the extended periods between blood feeding in ticks which can last for months or longer and their apparent inability to synthesize heme, it is reasonable to assume that ticks developed mechanisms to sequester and store heme critical to their normal function.

To date the only major hemolymph storage protein in ticks that has been sequenced and studied in detail is vitellogenin (Thompson *et al*., 2005, 2007). In ticks, Vg transcription occurs only in females and only after mating and feeding to repletion. The protein is synthesized by the fat body, midgut and to a lesser extent by the ovary of pre-ovipositing and ovipositing females (Thompson *et al*., 2007) and presumably secreted by these tissues into the hemolymph as vitellogenin. Deposition of vitellin into developing oocytes requires the appearance of the vitellogenin receptor in tick ovaries as recently determined by RNAi (Mitchell *et al*., 2007). Vitellogenin is a hemelipoglyco-carrier protein likely important in part because it transports heme from the female hemolymph to developing oocytes.

Chinzei (1983) was the first to show by polyacrylamide gel electrophoresis (PAGE), a major hemolymph protein in the soft tick, *Ornithodoros parkeri* Cooley. This protein was found in both male and female ticks throughout development from the larval to the adult stage (Guddera *et al.*, 2001, 2002a,b) and was characterized as a hemelipoglyco-carrier protein referred to as CP in *D. variabilis* (Guddera *et al.*, 2001, 2002a,b) or HeLp in *R. microplus* (Maya-Monteiro *et al.*, 2000, 2004).

Guderra *et al.* (2001) found that the molecular weight of CP in *D*. *variabilis* (DvCP) was 200 and 340 kDa by native PAGE and gel filtration chromatography, respectively, while Maya-Monteiro *et al.* (2000) reported an average molecular weight by native-PAGE and gel filtration of this protein from *R. microplus* (RmCP) as 354 kDa. In both organisms CP consists of two subunits, 92 and 98 kDa for DvCP and 92 and 103 kDa for RmCP. Reported N-terminal sequences obtained by Edman degradation for DvCP and RmCP (Guderra *et al*. 2001; Maya-Monteiro *et al*. 2000) suggested that the amino acid sequences were conserved among the Ixodida. CP was found to carry cholesterol, phospholipids, monoacylglycerides, triacylglycerides and free fatty acids in addition to heme (Guddera *et al*., 2001; Maya-Monteiro *et al.*, 2000). Maya-Monteiro *et al*. (2004) demonstrated that heme bound to RmCP had a reduced capacity to induce oxidative damage to phospholipid bilayers. Guddera *et al*. (2001) described a similar protein to CP in the soft tick, *Ornithodoros parkeri*, with an approximate molecular weight of 668 kDa. No protein sequence data for CP from soft ticks is yet available as well as any full nucleotide sequences from any tick.

In the current study, we report the first full length sequence of the major heme binding protein found in both male and female *D. variabilis* (CP) and the validation of its function by LC MS and studies of its developmental and tissue specific expression during the adult stage. We also examined the evolutionary relationship

of CP with other known storage proteins and its potential role in the development of tick hematophagy.

Results

DvCP sequence

The complete DvCP mRNA sequence, GenBank accession number DQ422963, is 4951 bp with a 5' untranslated region (UTR) from base pairs 1 to 43, a TGA stop codon at bp 4685 followed by the 3' UTR, an AATAAA polyadenylation signal and a 17-bp adenine tail. The 5' UTR is relatively short although similar in length to that of *D. variabilis* vitellogenin (DvVg; GB accession #AY885250). Comparison of the UTR's from DvCP and DvVg does not produce a significant alignment, however (13.4% identity). The 5' UTR of the lone star tick, *Amblyomma americanum*, HeLp (EF050790) herein referred to as AaCP, excluding the reported vector and SMART IIA oligo contamination, is 87% identical to that of DvCP (data not shown).

A 4641 bp open reading frame for DvCP encoding a protein precursor of 1547 amino acids from base pairs 44 to 4684 (Figure 1) was confirmed by tryptic digest LC tandem mass spectrometry (ABD83654). The protein precursor contains a 15 residue secretion signal predicted by the SignalP Server. An RXXR cleavage site common to storage proteins and vitellogenins, exists immediately upstream of the N-terminus of the 98 kDa subunit from amino acids 746 to 749. Tryptic digest LC-

MS/MS analysis confirmed that the 92 kDa subunit is residues 16 to 749 (734 amino acids total) and the 98 kDa subunit is residues 750 to 1547 (798 amino acids total).

Three motifs in the completed DvCP protein sequence were predicted by the Conserved Domains Search Tool—a lipoprotein n-terminal domain from residues 16 to 622, a domain of unknown function (DUF1943) from residues 667 to 945 characteristic of proteins with several large open beta sheets, and a von Willebrand factor type D domain at the carboxy-terminus from residues 1351 to 1522 (Figure 1). Initial tryptic digestion LC-MS/MS analysis indicated several glycosylated peptidic fragments and in fact six potential N-linked glycosylation sites are present in the DvCP sequence.

Comparison of DvCP with other tick sequences and non-tick polypeptides

Translations of EST data of putative CP from *A*. *variegatum*, *R*. *appendiculatus*, *R. microplus* and *I*. *scapularis* as well as *A*. *americanum* CP were aligned with DvCP (Figs. 2 and 3). The sequences shown in Figure 2 all contain a 15 residue secretion signal. Comparison of available sequence data reveals that RaTC1551 is 84.2 and 76.0% identical in the first 317 residues to DvCP and AaCP, respectively. The available 317 residues of the *R*. *appendiculatus* homolog RaTC1839 were found to be 65.6 and 69.1% identical to DvCP and AaCP, respectively. Additional ESTs from *R*. *microplus* and *I*. *scapularis* were also located in the Genbank database and aligned with DvCP (Figure 3).

A BLASTP for DvCP against the non-redundant Genbank database returns the best non-CP match (expect value: 3e-30) as vitellogenin (Vg-6 isoform a) from the nematode *Caenorhabditis elegans* (AACO4423). An unpublished 191 amino acid DvCP fragment from another laboratory (ABC94727) was previously misidentified as vitellogenin from tick ovaries as determined by BLASTP. The majority of proteins from BLASTP were vitellogenins and Vg-like proteins. A radial dendrogram was constructed by maximum likelihood analysis using *D*. *variabilis* and *A*. *americanum* CP sequences as well as Vg and Vg-like proteins from insects, crustaceans, mollusks, fish, chicken and crustaceans (Figure 4). DvCP and AaCP grouped with tick vitellogenin.

DvCP tissue specific and developmental expression

To study the mRNA levels of CP, a 909 bp region of the gene (bp 1279-2187) was used to produce a probe for Northern blot analysis. Expression of CP in adult female *D*. *variabilis* appears to be initiated by attachment to the host (and blood feeding) and increases with attachment duration (Figure 5). CP expression was also detected in adult male ticks that were attached for 5 d but was not determined for other time points. After attachment for 6 d, expression of CP in female ticks was greatest in the fat body and salivary gland compared to the levels found in the ovary and midgut (Figure 5).

Analysis of cement cone and saliva for the presence of DvCP

Since the CP message was found at high levels in the salivary gland, cement cone and saliva were collected from part-fed female *D*. *variabilis* and studied. Repeated collections of saliva and cement cone from *D*. *variabilis* analyzed by native-PAGE failed to show any band with similar mobility to that of hemolymph DvCP (Figure 6). To address whether DvCP exists in a modified form in cement cone, cement cone proteins were separated by sodium dodecyl sulfate (SDS)-PAGE and subjected to tryptic digestion LC-MS/MS. A positive control of cement cone spiked with hemolymph was used to ensure that DvCP could be located in a complex sample of proteins. Two protein molecular weight ranges from the SDS-PAGE were subjected to tryptic digestion LC-MS/MS--proteins ≥66 kDa and proteins <66 kDa. No DvCP fragments were identified in either sample while DvCP was found in the ≥66 kDa positive control (data not shown).

Discussion

Identity and Origin of DvCP

The cleavage site between the secretion signal and the mature N-terminus of DvCP (Figure 1) produces a mature protein sequence that is in agreement with the N-terminal sequence data for hemolymph DvCP previously published by Guddera *et al*. (2001) as F(V)EVGKEYV and *R. microplus* apoHeLp-B previously published by Maya-Monteiro *et al*. (2000) as FEVGKEYV. In addition in the current study, tryptic

digest MS fingerprinting of purified female adult hemolymph DvCP identified several fragments with the same amino acid sequences found in the conceptual translation of the DvCP message (Figure 1). Since the DvCP message had a low identity to the vitellogenin message from the American dog tick (Figure 3), was found in both male and female ticks (Figure 5) unlike the female specific Vg message (Thompson *et al*. 2007), and the increase in the DvCP message occurred in response to blood feeding (Figure 5, discussed in more detail later) and was not initiated by mating like that for the DvVg message (Thompson *et al*. 2007), this further argues that CP is not the egg yolk protein. Although the presence of the DvCP message was not examined in the larval and nymphal stages in the current study, Guddera *et al.* (2001) found the hemolyph DvCP protein resolved by electrophoresis was not only the predominant whole tick protein in male and female adults but also in nymphs and larvae of *D. variabilis*. Only trace amounts were found in eggs. Since the native protein in adult hemolymph contained lipids, carbohydrates, and heme and appeared to be the predominant protein in both males and females throughout development (except in eggs), Guddera *et al*. (2001) concluded that DvCP was the major storage protein in the hard tick, *D. variabilis*. Guddera *et al*. (2001) found an analogous protein in the soft tick, *O. parkeri*, suggesting that CP was a common feature of all ticks. Other possible functions of this protein will be discussed later.

The motifs in the completed DvCP protein sequence predicted by the Conserved Domains Search Tool (Marchler-Bauer and Bryant, 2004) are a

lipoprotein n-terminal domain that is a common attribute of proteins that bind lipids, carbohydrates and metals and further supports the putative storage function of this protein in ticks. Additional domains identified included DUF1943 (exact function unknown) that is found in proteins with several large open beta sheets and a von Willebrand factor type D domain (Jorieux *et al*., 2000) that is likely present to aid multimerization of the two CP subunits. The three domains identified in DvCP are also common attributes of vitellogenins. Although it is clear as already discussed that CP is not the major egg yolk protein, there are clearly functional similarities between CP and Vg in ticks including their ability to bind heme. The mechanism of heme-binding by DvVg or DvCP has not been determined.

Nene *et al*. (2004) reported several *R*. *appendiculatus* tentative consensus sequences, TC1551 and TC1839 which were homologous to the N-terminus of *R. microplus* apoHeLp-B (Maya-Monteiro *et al*., 2000). These were not completed CP sequences. However, the same cleavage site was predicated by SignalP in the putative translation of CP from AaCP and *A*. *variegatum* TC353 (Figure 2). The alignments of DvCP, AaCP and *A*. *variegatum*, *R*. *appendiculatus*, *R. microplus* and *Ixodes scapularis* homologs suggest that this protein is highly conserved among ixodid ticks. Nene *et al*. (2004) first demonstrated the similarity of AvTC353, RaTC1551 and RaTC1839 to the N-terminus data of apoHeLp-B (Maya-Monteiro *et al*., 2000). Nene *et al*. (2004) also identified *R*. *appendiculatus* clones from the downstream portion of the CP gene. We located similar ESTs from *R*. *microplus*

and *I*. *scapularis* (Figure 3). These alignments indicate that at least two forms of CP may exist in the same ixodid ticks; however all tryptic digestion LC-MS/MS data of hemolymph DvCP matched with our putative translation.

A phylogenetic analysis of DvCP, AaCP and other vitellogenins and Vg-like proteins showed that CP grouped most closely with tick Vgs (Figure 4). Babin *et al*. (1999) previously showed that apolipophorin II/I, apolipoprotein B, vitellogenin and microsomal triglyceride transfer proteins (MTP) are derived from a common ancestor and all belong to the large lipid transfer protein (LLTP) superfamily. Smolenaars *et al*. (2007) compared vertebrate and invertebrate sequences from the LLTP superfamily that all contained a large lipid transfer module or motif located toward the N-terminal and reported three major groups--apoB-like LLTPs, MTPs and Vg-like LLTPs. Based on the current analysis (Figure 4), DvCP appears to be most closely related to the vitellogenins.

Tissue-Specific Regulation of the DvCP message is different from DvVg

The electrophoretic analysis of whole body homogenates based on equal protein loading by Guddera *et al*. (2001) revealed that unfed females contained a greater abundance of DvCP than that of females after attachment for 10 d and replete females 14 d after detachment. DvCP levels in plasma however, based on equal volume loading, indicated that the protein was more abundant in 11 and 13 d part-

fed ticks than in replete ticks, 9-17 d after detachment. To better understand the regulation of CP during adult development, a 909 bp region of the CP message (bp 1279-2187) was used to probe Northern blots. From these studies it was apparent that the expression of CP in whole body of adult virgin females of *D. variabilis* was initiated after attachment to the host and with the initiation of blood feeding, and appeared to increase with host-attachment duration (Figure 5). CP expression was also detected in adult male ticks that were attached for 5 d at comparable levels to that found in the part fed, virgin females; CP levels were not determined for other time points in male development. Guddera *et al*. (2001) reported that blood feeding had no obvious effect on the abundance of CP in the plasma of adult male ticks.

After attachment for 6 d, expression of CP in female ticks was greatest in the fat body and salivary gland compared to the levels found in the ovary and midgut (Figure 5). These results appeared to be in agreement with the CP protein levels in the same tissues as reported previously. Guddera *et al*. (2002a) found CP protein was higher in salivary gland and fat body than that of ovary. Protein levels in the midgut were not determined.

The regulation of the DvVg message was recently studied in detail by Thompson *et al*. (2005). The appearance of the DvVg message is not initiated by blood feeding in virgin females as was the case for CP but by mating in part-fed females. Mating also increases the levels of 20-hydroxyecdysone and the injection of the same hormone into part-fed, virgin ticks attached to the host resulted in

increased levels of the Vg message and the production of vitellogenic eggs similar to that produced by replete, mated females. Apparently, even though DvCP and DvVg may share a common evolutionary origin and have similar functions in binding lipids, carbohydrates and heme, the whole body tissue sources and regulation of their messages differ. This further supports the conclusion discussed before that DvCP is not the yolk protein (Vg) but the major male and female storage protein important for development from the larval through the adult stage. It was interesting that blood feeding initiated the synthesis of the CP message in part-fed, virgin females (Figure 5) but Guddera et al. (2001) found that mating and blood feeding to repletion reduced the levels of CP protein and Thompson *et al*. (2005) found that mating and blood feeding to repletion increased the Vg message, Vg hemolymph protein and yolk deposition in the egg. Apparently, in the adult stage before the initiation of vitellogenesis, CP is the major tick storage protein. Then this function shifts to Vg in replete ticks as a result of mating and with increased levels of ecdysteroids. Additional studies are needed to confirm this hypothesis.

Possible Role of CP in Host Complementation

Previous studies, although contradictory, have suggested that CP protein might occur in tick saliva and in the cement cone and may be involved in host complementation. This view is partially supported in our current study by the fact

that one of the major sites for the CP message is the salivary gland (Figure 5). Shapiro *et al*. (1986) found a 94 kDa antigen in the cement cone of 3 and 5 d fed adult *R*. *appendiculatus*. In addition, Jaworski *et al*. (1992) showed that a 90 kDa antigen was present in salivary glands that cross-reacted with a 70 kDa cement antigen in *D*. *variabilis*, *A*. *americanum* and *R*. *sanguineus* but not a 90 kDa antigen. Jaworski et al. (1992) also could not find a 90 kDa protein by SDS-PAGE in the cement proteins from *D*. *variabilis* and *A*. *americanum*. Wang and Nuttall (1994) found a 98 kDa protein present in the salivary glands of unfed and in 2, 4, 6 and 8 d fed female *R*. *appendiculatus* and in saliva of 6 and 8 d fed female ticks of the same species. The 98 kDa protein in saliva appeared to be a dominant component of the saliva as well as the dominant hemolymph protein. Madden *et al*. (2002) analyzed the saliva of *A*. *americanum* and *A*. *maculatum* and found that the dominant protein in each sample had an identical N-terminal sequence to that of RaCP (Maya-Monteiro *et al*., 2000) and nearly identical to that of DvCP (Guddera *et al*., 2001).

In consideration of the conflicting literature evidence that CP might be important in host complementation and our own work that showed that CP message was found in salivary gland, repeated attempts were made to resolve CP from saliva by native-PAGE and from cement cones by SDS-PAGE from adults of the American dog tick. Because of the high levels of CP in hemolymph and considering that injections into the hemolymph of dopamine and pilocarpine were required to initiate salivation, special attention was given to prevent any contamination of the saliva by

hemolymph. No bands were found with a similar mobility to that of hemolymph DvCP in any of these studies (Figure 6). To address whether DvCP exists in a modified form in cement cone, a possibility suggested by the results of a cross reactive serum in studies by Jaworski *et al*. (1992), cement cone proteins were separated by SDS-PAGE and the proteins ≥66 kDa and <66 kDa analyzed by LC MS/MS. A positive control was constructed with cement cone spiked with hemolymph containing DvCP. Again in these studies, no DvCP fragments could be found in the cement cone for either molecular weight range tested while DvCP was identified in the positive control sample of proteins ≥66 kDa (data not shown). These results suggest that DvCP is not a component of tick saliva and cement cone at the detection limit of silver staining or below the 50-75 ng detection limit of LC-MS/MS (for an ~100 kDa protein for the latter).

The role of the CP message in the salivary gland (Figure 5) as well as CP protein (Guddera et al., 2002a) is not clear but suggests that the salivary gland might be a source for the hemolymph CP. Alternatively, the lack of CP in the saliva might result from the artificial method used to collect the saliva. However, no CP was found in the cement cone as had been suggested by others. It would also be surprising that CP which is the major heme, glycolipo-storage protein in ticks and the predominant tick protein throughout development is also a critical component in host complementation. Further studies will be needed to resolve these questions. In our

exhaustive attempts to find CP in saliva and in the cement cone of *D. variabilis* using the methods as described, the protein could not be found.

Role of CP in heme sequestration and evolution to blood feeding

Bloodmeal digestion in ticks is initiated in the midgut (Coons *et al*., 1986; Tarnowski and Coons, 1989) with the lysis of host erythrocytes presumably to release serine proteases (Miyoshi *et al*., 2007). Uptake of the bloodmeal components by the digest cells in the midgut is facilitated by receptor mediated endocytosis via fluid-phase endocytosis and clathrin-coated pits (Coons *et al*., 1986). Intracellular digestion by heterophagy, the fusion of primary lysosomes and endosomes to form secondary lysosomes, results in the breakdown of host hemoglobin. The resulting heme becomes concentrated into dense residual bodies (Tarnowski and Coons, 1989) or hemosomes (Lara *et al*., 2003) which serve as a hematin sequestration mechanism to prevent reduction-oxidation reactions from forming harmful free radicals. During intracellular digestion, the digest cells are found along the basal lamina where heme is transferred to the hemocoel. The exact transfer mechanism is not yet clear but preliminary data suggest that a membrane-protein transporter is involved (Lara *et al*., 2005). Apparently, heme is then sequestered in the hemolymph by CP in a 2:1 heme to CP molar ratio (Maya-Monteiro *et al*., 2000). During vitellogenesis heme is also transferred to vitellogenin

directly and/or via CP (Thompson *et al*., 2007) with Vg serving as the major route of heme trafficking into the developing oocytes.

Both CP and Vg appear to be critical components for the sequestration and utilization of host heme in the absence of *de novo* heme synthesis in the ixodidae and for the extended periods of non-feeding that occurs in ticks, sometimes for months at a time. As a result, these two proteins appear to be critical to the evolution of hematophagy.

In summary, CP in ixodid ticks is a non-sex linked highly conserved hemelipoglycoprotein that appears to be derived from an ancestral Vg-like gene. Of the Vg and Vg-like proteins sequenced thus far, the most closely related sequence is that of a vitellogenin from *C*. *elegans*. CP appears to be ubiquitous in all life stages of ticks. Its expression in female adults is initiated by attachment to the host and blood feeding. In addition to serving as a storage protein for cholesterol, phospholipids, monoacylglycerides, triacylglycerides and free fatty acids (Guddera *et al*., 2001), CP binds heme for utilization by the tick and as a sequestration mechanism to prevent the formation of potentially damaging radical species (Maya-Monteiro *et al*., 2004). DvCP was not detected in the saliva or cement in *D*. *variabilis* despite reports of its presence in high levels in saliva (Madden *et al*., 2002). The functional significance of high levels of CP mRNA and protein in the salivary gland needs further study.

Experimental Procedures

Tick Rearing

The American dog tick, *D*. *variabilis*, was reared as previously described (Sonenshine, 1993) from specimens originally collected near Righmond, VA. Adult ticks were confined within plastic capsules attached to New Zealand White Rabbits, *Oryctolagus cuniculus,* and allowed to feed as required. Larvae and nymphs were fed on Norway rats, *Rattus norvegicus*. Rearing conditions were 26 ± 1ºC, 92 ± 6% relative humidity and 14:10 (L:D). .

Tissue, Saliva and Cement Cone Collection

Hemolymph was collected from 5-6 d partially fed virgin female ticks as described by Johns *et al*. (1998). Briefly, ticks were mounted ventral side up on glass slides with double-sided tape and one of the forelegs was amputated at the coxal-trochanteral joint. Pressure was applied to the body of the tick to recover hemolymph, which was collected with a glass micropipette and then stored at -80°C until further use. Fat body, midgut, ovary and salivary gland were dissected from 6 d partially fed virgin females, washed in ice-cold phosphate-buffered saline (PBS; Dulbecco's phosphate buffer solution, pH 7.4, 8 mM sodium phosphate, 2 mM potassium phosphate, 0.14 M sodium chloride, 0.01 M potassium chloride; Pierce, Rockford, IL) and stored in Tri-Reagent (Sigma Chemical Co., St. Louis, MO) at -80°C until further use.

Saliva was collected from 5-6 d part fed virgin female ticks by injecting 2-3 µl of a solution containing 0.1% dopamine, 150 mM sodium chloride, 2 mM calcium chloride, 2.5% pilocarpine and 0.02% DMSO into the hemocoel through the alloscutum with a 10 µl Hamilton syringe (Hamilton, Reno, NV) fitted with a 30 gauge needle. Cement cones (approximately 10 per sample) were collected from 5-6 d fed virgin female ticks, and the cement was homogenized in a 1.5 ml centrifuge tube with a plastic pestle (USA Scientific, Ocala, FL) in 100 µl of PBS. Saliva and homogenized cement cone samples were stored at –80ºC until further use.

Electrophoresis

To purify DvCP for *de novo* sequencing by tryptic digestion liquid chromatography tandem mass spectroscopy (LC-MS/MS; described below), hemolymph was diluted 1:100 in PBS and then 6 µl was diluted with native sample buffer (Invitrogen, Carlsbad, CA). Proteins were separated electrophoretically with an Invitrogen Xcell SureLock™ electrophoresis apparatus with an 8-16% tris-glycine polyacrylamide gel using appropriate buffers from Invitrogen for native-PAGE. Electrophoresis was conducted at 134 V for 4 h at 4°C. The gel was stained for 1 h with Coomassie Brilliant Blue R-250 (45 ml methanol, 10 ml glacial acetic acid, 45 ml ddH_2O, 0.25 g Coomassie Brilliant Blue R-250 (Fisher Scientific, Pittsburgh, PA)) and then destained overnight with a mixture of 30% methanol and 10% glacial acetic acid. A

band at ~200 kDa compared to native-PAGE molecular weight markers (Sigma, St. Louis, MO) was excised for *de novo* sequencing by LC-MS/MS.

For the direct comparison of hemolymph and cement cone and preparation for tryptic digestion mass fingerprinting, samples were analyzed by SDS-PAGE. Hemolymph was diluted 1:10 in PBS and 1 µl was added to 1 µl of 100 mM dithiothreitol (DTT; Fisher Scientific), 5 µl SDS buffer (Invitrogen), and 3 µl ddH_2O. The sample was incubated at 100°C for 5 min and then briefly centrifuged before loading the entire contents on the gel. Cement cone samples were prepared with 5 µl of cement cone homogenate, 2 µl DTT, 10 µl SDS buffer and 3 µl ddH_2O. The sample was incubated at 100°C for 5 min and then centrifuged for 5 min at 12,000 x *g* at room temperature and 10 µl of the resulting supernatant was used for electrophoresis. A third sample used as a positive control for LC-MS/MS consisted of 4 µl of cement cone homogenate, 6 µl of hemolymph diluted 1:100 in PBS, 2 µl DTT and 10 µl of SDS buffer. The sample was incubated and centrifuged as mentioned above and then 12 µl of supernatant was used for electrophoresis. Benchmark pre-stained protein ladder (Invitrogen) was used for estimation of molecular weight. Proteins were separated with a 4-12% tris-glycine gel (Invitrogen) at 125 V for 90 min at room temperature. The gels were either stained with Coomassie Brilliant Blue R-250 (described earlier) for tryptic-digestion mass fingerprinting compatibility or silver stained with standard reagents from Bio-Rad

(Bio-Rad Laboratories, Hercules, CA) according to the manufacturer's recommendations. The experiments were replicate three times.

For the direct comparison of hemolymph and saliva, samples were analyzed by native-PAGE. Hemolymph samples for native-PAGE were diluted 1:10 in PBS and 2 µl were added to 3 µl ddH_2O and 5 µl of native sample buffer (Invitrogen). Saliva samples consisted of 6 µl saliva, 4 µl ddH_2O and 10 µl native sample buffer. The hemolymph and saliva samples were separated with a 4-12% tris-glycine gel (Invitrogen) at 122 V for approximately 4 h at 4°C. NativeMark (Invitrogen) molecular weight marker was used to estimate protein molecular weight. The gels were silver stained as mentioned earlier. The experiment was replicated three times.

Digestion-mass fingerprinting

Hemolymph DvCP was gel purified by native-PAGE in preparation for *de novo* sequencing by tryptic digestion LC-MS/MS (described earlier) and submitted to the W.M. Keck Biomedical Mass Spectrometry Laboratory, Biomolecular Research Facility, University of Virginia Health System, Charlottesville, VA.

All other tryptic digestion mass fingerprinting samples (described earlier) were analyzed at the North Carolina State University Metabolomics and Proteomics Laboratory, Raleigh, NC. Methods for in-gel tryptic digestion and elution of the peptide fragments have been previously described (Jiménez *et al*. 1996;

Shevchenko *et al*. 2006). The 92 and 98-kDa subunits of DvCP were excised and prepared separately from SDS-PAGE gels (described earlier) for confirmation of the putative DvCP amino acid sequence. To assess whether DvCP was present in the cement cone the following SDS-PAGE purified samples were analyzed: combined hemolymph DvCP 92 and 98-kDa subunits; all cement cone peptides ≥66 kDa; all cement cone peptides <66 kDa; cement cone and hemolymph peptides ≥66 kDa (the positive control). Five microliter samples were analyzed on a Thermo LTQ linear ion trap mass spectrometer with an IonMax electrospray ionization source interfaced to a Phenomenex Jupiter 4 µm Proteo C18 reverse phase column. Peptides were eluted with a linear buffer gradient in 30 min with a 28 µl/min flow rate. The linear buffer gradient was as follows: 3% buffer A/97% buffer B to 40% buffer A/60% buffer B (buffer A - 0.01% TFA, 0.005% heptaflurobutyric acid (HFBA) in ddH_2O; buffer B - 0.01% TFA, 0.005% HFBA in acetonitrile). The ion source was operated at 4.5 kV.

The peptides were analyzed as a big 4 experiment, utilizing Dynamic Exculsion (C/R). As such, a survey scan (MS-only mode) was followed by four rounds of data-dependent MS/MS scans on the four most abundant ions in the survey scan. Ions chosen for MS/MS were then placed onto an exclusion list for a total of 180 s. The approach allows for the detection and characterization of minor components within the sample.

Database searches were made in Bioworks Browser 3.1 (Thermo, Waltham, MA). The initial database searches were made against a custom local database file that contained the *D*. *variabilis* hemelipoglycoprotein amino acid sequence (Accession #ABD83654) and UniprotKB-TrEMBL (www.pir.uniprot.org; download date, 21 September 2006).

RNA isolation

Total RNA was isolated from whole tick bodies (unfed female, 6 h fed female, 24 h fed female, 5 d fed male and 6 d fed female) or dissected tick tissues (described above) using Tri-Reagent according to the manufacturer's recommendations. RNA pellets were dissolved in 10 μM aurintricarboxylic acid to prevent degradation (Hallick *et al*. 1977). The purity and concentration of each sample was measured with a Spectromax 384 Plus spectrophotometer (Molecular Devices Corporation, Sunnyvale, CA), and then the samples were stored at -80°C until further use.

Cloning and Sequencing of CP

Two micrograms of 6 d fed female whole body total RNA was reverse transcribed with Powerscript reverse transcriptase (Clontech, Mountain View, CA) and primed with Oligo-dT/Ankr (5'-CCCACAGGCACTACGATGTA$(T)_{17}$V-3'). First-strand cDNA was PCR amplified with forward primer CPFP11 (5'-CCTKCTCTTYGTCCAGACCGCTC-3') and reverse primer Ankr (5'-

CCCACAGGCACTACGATGTA-3') using *Taq* DNA Polymerase (New England Biolabs, Ipswich, MA) according the manufacturer's recommendations. Thermal cycling was carried out at 95ºC for 2 min, 35 cycles of 95ºC for 30 s, 61ºC for 30 s, and 72ºC for 1 min, followed by a final extension of 72ºC for 5 min. The resulting 422 bp fragment from each of the 3 PCR amplifications was submitted to the Nucleic Acids Research Facilities at Virginia Commonwealth University for sequencing.

Reverse primer CPRP4 (5'-GTGAATGCGCGAGGAACGGT-3') and forward primer CPFP6 (5'-TTCGAGGTCGGCAARGARTA-3') located 42 bp downstream of the start codon, using Phusion^{TM} high fidelity DNA polymerase (Finnzymes, Espoo, Finland), was used to PCR amplify a 4712 bp region of the CP gene. Initially CPFP6 and CPRP4 were used to obtain sequence data; additional primers were then designed to obtain the complete sequence of the 4712 bp region. Each fragment was sequenced a minimum of 3 times.

The upstream region of the gene was completed by 5' RACE. Briefly, cDNA was reverse transcribed with Powerscript reverse transcriptase and primed with CPRP5 (5'-ACCGTGAAGACCGTCCTCCT-3'). Included in the reaction was the SMART IV^{TM} oligonucleotide (Clontech) which due to the terminal transferase activity of the enzyme, binds to the $C_{(3)}$ underhang and switches templates during transcription. The resulting cDNA was PCR amplified with 5' PCR primer (Clontech) and CPRP5. Thermal cycling was carried out at 95ºC for 5 min, 35 cycles of 95ºC for 30 s, 63ºC for 30 s, 72ºC for 75 s, and a final extension of 72ºC for 5 min.

Northern blot analysis

Total RNA was isolated and quantitated as indicated previously. Aliquots of total RNA were denatured by glyoxal treatment, and 5 µg of each sample was separated by electrophoresis in a 1.25% agarose gel according to the phosphate protocol of Sambrook and Russell (2001). Approximately 100 ng of glyoxylated digoxigenin-labeled molecular weight marker II (Roche Diagnostics GmbH, Mannheim, Germany) was included. RNA was transferred by capillary action to a nylon membrane (Roche) followed by UV crosslinking according to the manufacturer's recommendations. CP and β-actin probes were digoxigenin-labeled by PCR amplification. The CP probe was PCR amplified from a 909 bp region of the gene (Accession #DQ422963) from base pairs 1279 to 2187 with the inclusion of digoxigenin-11-dUTP according to the manufacturers recommendations (Roche Diagnostics GmbH, Mannhein, Germany). A β-actin (Genbank Accession #EF488512) probe was also PCR amplified from a 300 bp region of the gene and used to ensure equal loading of RNA. Blots were pre-hybridized, hybridized and washed according to the manufacturer's recommendation, except that all elevated temperature manipulations were performed at 50ºC instead of 68ºC. Chromogenic detection was performed using NBT/BCIP according to the manufacturer's recommendations.

Bioinformatics

The location of the putative CP secretion signal was determined with the SignalP prediction tool (Bendtsen *et al*., 2004). Domains were searched with the Conserved Domains Search tool (Marchler-Bauer and Bryant, 2004) available on the Genbank website (www.ncbi.nlm.nih.gov). All alignments were conducted with ClustalW (www.ebi.ac.uk/clustalw). Contiguous sequences of AvCP and RaCP were produced from EST data available at The Institute for Genomic Research (www.tigr.org). An alignment of proteins identified by Blastp (Altschul *et al*., 1997) was uploaded to the MultiPhyl website (http://distributed.cs.nuim.ie/multiphyl.php) for maximum likelihood analysis. Twenty-eight protein sequences from 27 taxa were selected to compare DvCP with Vg and Vg-like proteins from the groups represented in the blast result (ticks, insects, crustaceans, nematodes, mollusks, fish and chicken). Using the AIC1 model, trees were searched with the subtree pruning and regrafting algorithm with an initial neighbor joining tree. Bootstrapping was performed with 100 replicates.

Acknowledgments

This project was supported by a Grant (IBN-0315179) from the National Science Foundation and research support to R.M.R. from the North Carolina Agricultural Research Service. We thank Dr. Charles A. Apperson, Dr. Joanna A. Miller and Dr. Christina M. Grozinger at North Carolina State University for thoughtful discussions and review of the manuscript, Dr. Nigel Deighton at the North Carolina State

University Metabolomics and Proteomics Laboratory for assistance with LC-MS/MS analyses and Callie Barber for gel imaging.

References

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

Babin, P.J., Bogerd, J., Kooiman, F.P., Van Marrewijk, W.J.A. and Van der Horst, D.J. (1999) Apolipophorin II/I, apolipoprotein B, vitellogenin, and microsomal triglyceride transfer protein genes are derived from a common ancestor. *J Mol Evol* **49**: 150-160.

Bendtsen, J.D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol **340**: 783-795.

Braz, G.R.C., Coelho, H.S.L., Masuda, H. and Oliveira, P.L. (1999) A missing metabolic pathway in the cattle tick *Boophilus microplus*. *Curr Biol* **9**: 703-706.

Chinzei, Y. (1983) Quantitative changes of vitellogenin and vitellin in adult female ticks, *Ornithodoros moubata*, during vitellogenesis. *Mie Med J* **27**: 117-127.

Coons, L.B., Rosell-Davis, R. and Tarnowski, B.I. (1986) Bloodmeal digestion in ticks. In *Morphology, Physiology and Behavioral Biology of Ticks*, pp. 248-279. (Sauer, J.R. and Hair, J.A. eds). Ellis Horwood, Chinchester, England.

Furuyama, K., Kaneko, K. and Vargas, V. (2007) Heme as a magnificent molecule with multiple missions: heme determines its own fate and governs cellular homeostasis. *Tohoku J. Exp. Med* **213**: 1-16.

Guddera, N.P., Neese, P.A., Sonenshine, D.E., Apperson, C.S. and Roe, R.M. (2001) Developmental profile, isolation, and biochemical characterization of a novel lipoglycoheme-carrier protein from the American dog tick, *Dermacentor variabilis* (Acari: Ixodidae) and observations on a similar protein in the soft tick, *Ornithodoros parkeri* (Acari: Argasidae). *Insect Biochem Molec Biol* **31**: 299-311.

Guddera, N.P, Sonenshine, D.E., Apperson, C.S. and Roe, R.M. (2002a) Tissue distribution and characterization of predominant hemolymph carrier proteins from *Dermacentor variabilis* and *Ornithodoros parkeri*. *J Insect Physiol* **48**: 161-170.

Guddera, N.P, Sonenshine, D.E., Apperson, C.S. and Roe, R.M. (2002b) Hemolymph proteins in ticks. *J Insect Physiol* **48**: 269-278.

Hallick, R.B, Chelm, B.K., Gray, P.W. and Orozco, Jr., E.M. (1977) Use of aurintricarboxylic acid as an inhibitor of nucleases during nucleic acid isolation. *Nucleic Acids Res* **4**: 3055-3064.

Hamza, I., Chauhan, S., Hassett, R. and O'Brian, M.R. (1998) The bacterial IRR protein is required for coordination of heme biosynthesis with iron availability. *J Biol Chem* **273**: 21669-21674.

Jaworski, D.C., Rosell, R., Coons, L.B. and Needham, G.R. (1992) Tick (Acari: Ixodidae) attachment cement and salivary gland cells contain similar immunoreactive polypeptides. *J Med Entomol* **29**: 305-309.

Jiménez, C.R., Huang, L., Qiu Y. and Burlingame, A.L. (1998) In-gel digestion of proteins for MALDI-MS fingerprint mapping. In *Current Protocols in Protein Science* (Coligan, J.E. ed), pp. 16.4.1-16.4.5. John Wiley & Sons, Inc., Brooklyn, NY, USA.

Johns, R., Sonenshine, D.E. and Hynes, W.L. (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.

Jorieux, S., Fressinaud, E., Goudemand, J., Gaucher, C., Meyer, D. and Mazurier, C. (2000) Conformational changes in the D' domain of von Willebrand factor induced by CYS 25 and CYS 95 mutations lead to factor VIII binding defect and multimeric impairment. *Blood* **95**: 3139-3145.

Lara, F.A., Lins, U., Bechara, G.H. and Oliveira, P.L. (2005) Tracing heme in a living cell: hemoglobin degradation and heme traffic in digest cells of the cattle tick *Boophilus microplus*. *J Exp Biol* **208**: 3093-3101.

Lara, F.A., Lins, U., Paiva-Silva, G., Almeida, I.C., Braga, C.M., Miguens, F.C., Oliveira, P.L. and Dansa-Petretski, M. (2003) A new intracellular pathway of heme detoxification in the midgut of the cattle tick *Boophilus microplus*:

aggregation inside a specialized organelle, the hemosome. *J Exp Biol* **206**: 1707-1715.

Law, J.H., Ribeiro, J.M.C. and Wells, M.A. (1992) Biochemical insights derived from insect diversity. *Ann Rev Biochem* **61**: 87-111.

Machado, E.A., Oliveira, P.L., Moreira, M.F., de Souza, W. and Masuda, H. (1998) Uptake of *Rhodnius* heme-binding protein (RHBP) by the ovary of *Rhodnius prolixus*. *Arch Insect Biochem Physiol* **39**: 133-143.

Madden, R.D., Sauer, J.R. and Dillwith, J.W. (2002) A proteomics approach to characterizing tick salivary secretions. *Exp Appl Acarol* **28**: 77-87.

Mans, B.J. and Neitz, A.W.H. (2004) Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. *Insect Biochem Mol Biol* **34**: 1-17.00000

Marchler-Bauer, A. and Bryant, S.H. (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* **32**: W327-31.

Maya-Monteiro, C.M., Alves, L.R., Pinhal, N., Abdalla, D.S.P. and Oliveira, P.L. (2004) HeLp, a heme-transporting lipoprotein with an antioxidant role. *Insect Biochem Mol Biol* **34**: 81-87.

Maya-Monteiro, C.M., Daffre, S., Logullo, C. Lara, F.A., Alves, E.W., Capurro, M.L., Zingali, R., Almeida, I.C. and Oliveira, P.L. (2000) HeLp, a heme lipoprotein from the hemolymph of the cattle tick, *Boophilus microplus*. *J Biol Chem* **275**: 36584-36589.

Mitchell, R.M., Ross, E., Osgood, C., Sonenshine, D.E., Donohue, K.V., Khalil, S.M., Thompson, D.M. and Roe, R.M. (2007) Molecular characterization, tissue-specific and RNAi knockdown of the first vitellogenin receptor from a tick. *Insect Biochem Mol Biol* **37**: 375-388.

Miyoshi, T., Tsuji, N., Islam, M.K., Huang, X., Motobu, M., Alim, M.A. and Fujisaki, K. (2007) Molecular and reverse genetic characterization of serine proteinase-induced hemolysis in the midgut of the ixodid tick *Haemaphysalis longicornis*. *J Insect Physiol* **53**: 195-203.

Nene, V., Lee, D., Kang'a, S., Skilton, R., Shah, T., de Villiers, E., Mwaura, S., Taylor, D., Quackenbush, J. and Bishop, R. (2004) Genes transcribed in the salivary glands of female *Rhipicephalus appendiculatus* ticks infected with *Theileria parva*. *Insect Biochem Mol Biol* **34**: 1117-1128.

Oliveira, P.L., Kawooya, J., Ribeiro, J.M.C., Meyer, T., Poormam, R., Alves, E.W., Walker, F.A., Machado, E.A, Nussenzveig, R.H., Padovan, G.J. and Masuda, H. (1995) A heme-binding protein from hemolymph and oocytes of the blood-sucking insect, *Rhodnius prolixus*. *J Biol Chem* **270**: 10897-10901.

Sambrook, J. and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shapiro, S.Z., Voigt, W.P. and Fujisaki, K. (1986) Tick antigens recognized by serum from a guinea pig resistant to infestation with the tick *Rhipicephalus appendiculatus*. *J Parasitol* **72**: 454-463.

Shevchenko, A., Tomas, H., Havliš, J., Olsen, J.V. and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protocols* **1**: 2856-2860.

Smolenaars, M.M.W., Madsen, O., Rodenburg, K.W. and Van der Horst, D.J. (2007) Molecular diversity and evolution of the large lipid transfer protein superfamily. *J. Lipid Res* **48**: 489-502.

Sonenshine, D.E. (1993) *Biology of Ticks*, Vol. 1. Oxford University Press, New York.

Tarnowski, B.I. and Coons, L.B. (1989) Ultrastructure of the midgut and bloodmeal disgestion in the adult tick *Dermacentor variabilis*. *Exp Appl Acarol* **6**: 263-289.

Thompson, D.M., Khalil, S.M.S., Jeffers, L.A., Ananthapadmanaban, U., Sonenshine, D.E., Mitchell, R.D., Osgood, C.J., Apperson, C.S. and Roe, R.M. (2005) In vivo role of 20-hydroxyecdysone in the American dog tick, *Dermacentor variabilis* (Say). *J Insect Physiol* **51**: 1105-1116.

Thompson, D.M., Khalil, S.M.S., Jeffers, L.A., Sonenshine, D.E., Mitchell, R.D., Osgood, C.J. and Roe, R.M. (2007) Sequence and the developmental and tissue-specific regulation of the first complete vitellogenin messenger RNA from ticks responsible for heme sequestration. *Insect Biochem Molec Biol* **37**: 363-374.

Wang, H. and Nuttall, P.A. (1994) Comparison of the proteins in salivary glands, slaiva and haemolymph of *Rhipicephalus appendiculatus* female ticks during feeding. *Parasitology* **109**: 517-523.

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

Figure Captions

Figure 1. A. Putative amino acid sequence of *Dermacentor variabilis* hemelipoglycoprotein (DvCP; ABD83654). Amino acids in bold indicate *de novo* sequenced fragments obtained by tryptic digestion mass fingerprinting. Predicted secretion signal shown with single underline (bps 1-15), lipoprotein n-terminal domain shown with wavy underline (16-622), domain of unknown function DUF1943 shown with dashed underline (667-945), RXXR cleavage site shown with box (746-749), and von Willebrand factor type D domain shown with double underline (1351-1522). Six potential N-linked glycosylation sites with the NX(S/T) motif are shown with grey background. B. Graphical representation of the key components of the DvCP message.

Figure 2. ClustalW alignment of the conceptual translation of the first 400 amino acids from the N-termini (including the predicted secretion signal from residues 1-15) of CP from *Dermacentor variabilis* CP (DvCP), *Amblyomma americanum* CP (AaCP), as well as tentative consensus sequences from *Amblyomma veriegatum* (AvTC353) and *Rhiphicephalus appendiculatus* (RaTC1551 and RaTC1839). Light grey background with black text represents identical residues, dark grey/white text is conservative residues and dark grey/black text is similar residues.

Figure 3. ClustalW alignment of the conceptual translations of *Dermacentor variabilis* (DvCP) from residues 961 to the carboxy terminus and *Amblyomma americanum* CP, as well as tentative consensus sequences from *Rhipicephalus (Boophilus) microplus* (RmTC172), *Rhipicephalus appendiculatus* (RaTC1299) and *Ixodes scapularis* (IsTC8328 and IsTC16). Light grey background with black text represents identical residues, dark grey/white text is conservative residues and dark grey/black text is similar residues.

Figure 4. Unrooted phylogenetic tree of *Dermacentor variabilis* CP with vitellogenin (Vg) from insects, crustaceans, *Caenorhabditis elegans*, two crustacean clotting proteins and *Amblyomma americanum* CP. All proteins were chosen from a BLASTp result of DvCP. Maximum likelihood analysis was used to construct the tree; numbers at each node indicate bootstrap values. Accession numbers are *Aedes aegypti* Vg AAA18221, *Amblyomma americanum* CP ABK40086, *Apis mellifera* Vg NP_001011578, *Blatella germanica* Vg CAA06379, *Bombyx mori* BAA06397, *Caenorhabditis elegans* Vg BAB69831, *Cherax quadricarinatus* Vg AAG17936, *Crassostrea gigas* Vg BAC22716, *Danio rerio* Vg NP_739573, *Danio rerio* Vg-like protein CAK03619, *Daphia magna* Vg-Superoxide Dismustase BAE94323, *Dermacentor variabilis* CP ABD83654, *Dermacentor variabilis* Vg1 AAW78557, *Fundulus heteroclitus* Vg AAA93123, *Gallus gallus* Vg BAA13973, *Gambusia affinis* Vg BAD93698, *Litopenaeus vannamei* Vg AAP76571,

Macrobrachium rosenbergii Vg BAB69831, *Melanogrammus aeglefinus* Vg AAK151157, *Pacifastacus leniusculus* clotting protein AAD16454, *Penaeus monodon* clotting protein AAF19002, *Periplaneta americanum* Vg2 Q9BPS0, *Pimpla nipponica* AAC32024, *Rhipicephalus (Boophilus) microplus* Vg AAA92143, *Riptortus clavatus* Vg AAB72001, *Samia cynthia ricini* Vg BAB32641, *Saturnia japonica* Vg BAD91195, *Tenebrio molitor* melanization protein BAB03250 and *Toxorhynchites amboinensis* Vg-C AAV31932.

Figure 5. Tissue specific and developmental expression of *Dermacentor variabilis* hemelipoglycoprotein determined by Northern blot analysis. Times shown refer to the duration of attachment to the host. MG – midgut; FB – fat body; OV – ovary; SG – salivary gland. The size in nucleotides is shown for the molecular weight markers in lane 1.

Figure 6. Comparison of hemolymph proteins with cement cone or saliva from blood-fed female *Dermacentor variabilis*. A. SDS-PAGE of Lane 1, molecular weight markers; Lane 2, hemolymph; Lane 3, cement cone. B. Native-PAGE of Lane 1, molecular weight markers; Lane 2, hemolymph; Lane 3, saliva. Proteins were visualized by silver stain. Asterisks indicate the location of DvCP subunits in A and native DvCP in B.

Figure 1.

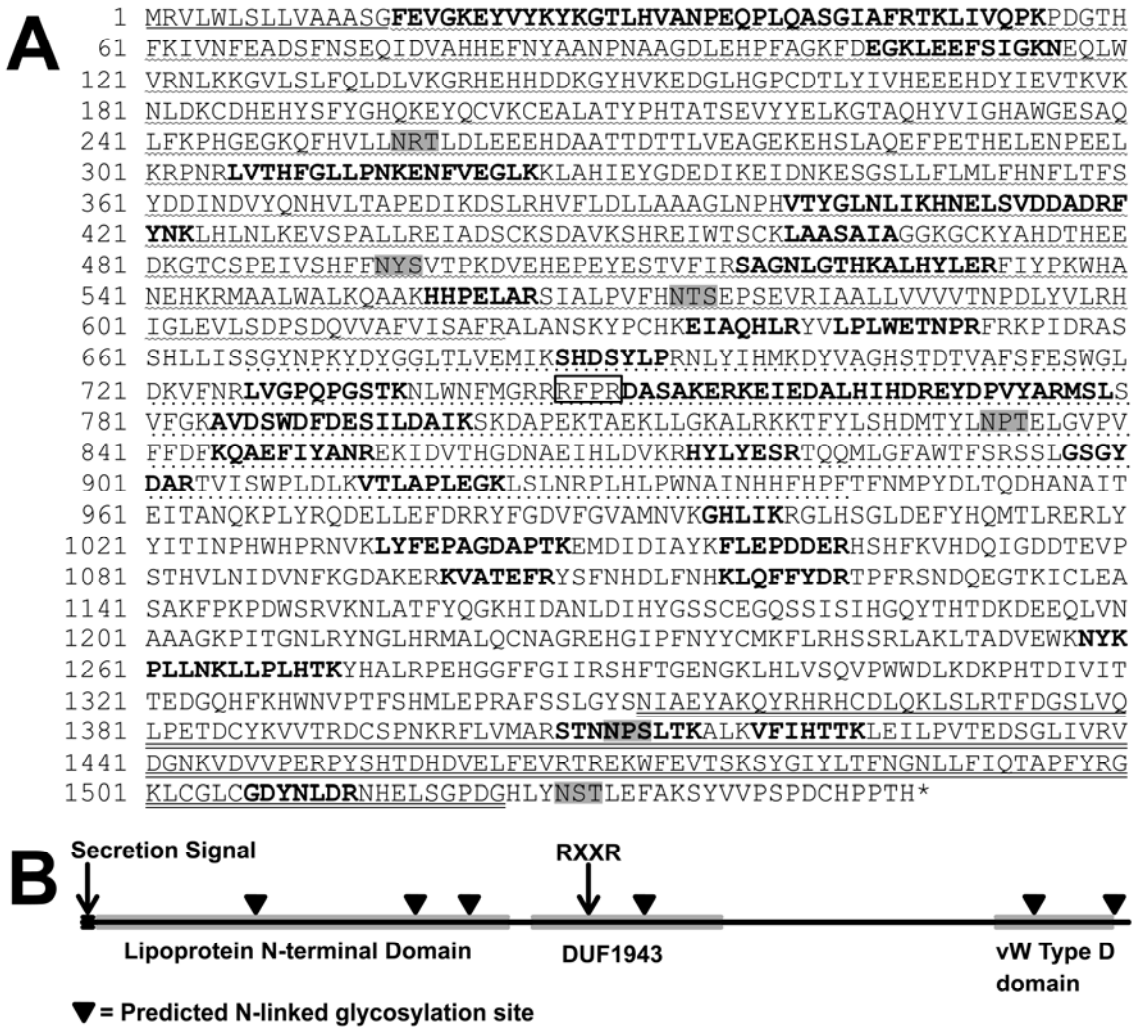

A
1 MRVLWLSLLVAAASGFEVGKEYVYKYKGTLHVANPEQPLQASGIAFRTKLIVQPKPDGTH
61 FKIVNFEADSFNSEQIDVAHHEFNYAANPNAAGDLEHPFAGKFDEGKLEEFSIGKNEQLW
121 VRNLKKGVLSLFQLDLVKGRHEHHDDKGYHVKEDGLHGPCDTLYIVHEEEHDYIEVTKVK
181 NLDKCDHEHYSFYGHQKEYQCVKCEALATYPHTATSEVYYELKGTAQHYVIGHAWGESAQ
241 LFKPHGEGKQFHVLLNRTLDLEEEHDAATTDTTLVEAGEKEHSLAQEFPETHELENPEEL
301 KRPNRLVTHFGLLPNKENFVEGLKKLAHIEYGDEDIKEIDNKESGSLLFLMLFHNFLTFS
361 YDDINDVYQNHVLTAPEDIKDSLRHVFLDLLAAAGLNPHVTYGLNLIKHNELSVDDADRF
421 YNKLHLNLKEVSPALLREIADSCKSDAVKSHREIWTSCKLAASAIAGGKGCKYAHDTHEE
481 DKGTCSPEIVSHFFNYSVTPKDVEHEPEYESTVFIRSAGNLGTHKALHYLERFIYPKWHA
541 NEHKRMAALWALKQAAKHHPELARSIALPVFHNTSEPSEVRIAALLVVVVTNPDLYVLRH
601 IGLEVLSDPSDQVVAFVISAFRALANSKYPCHKEIAQHLRYVLPLWETNPRFRKPIDRAS
661 SHLLISSGYNPKYDYGGLTLVEMIKSHDSYLPRNLYIHMKDYVAGHSTDTVAFSFESWGL
721 DKVFNRLVGPQPGSTKNLWNFMGRRRFPRDASAKERKEIEDALHIHDREYDPVYARMSLS
781 VFGKAVDSWDFDESILDAIKSKDAPEKTAEKLLGKALRKKTFYLSHDMTYLNPTELGVPV
841 FFDFKQAEFIYANREKIDVTHGDNAEIHLDVKRHYLYESRTQQMLGFAWTFSRSSLGSGY
901 DARTVISWPLDLKVTLAPLEGKLSLNRPLHLPWNAINHHFHPFTFNMPYDLTQDHANAIT
961 EITANQKPLYRQDELLEFDRRYFGDVFGVAMNVKGHLIKRGLHSGLDEFYHQMTLRERLY
1021 YITINPHWHPRNVKLYFEPAGDAPTKEMDIDIAYKFLEPDDERHSHFKVHDQIGDDTEVP
1081 STHVLNIDVNFKGDAKERKVATEFRYSFNHDLFNHKLQFFYDRTPFRSNDQEGTKICLEA
1141 SAKFPKPDWSRVKNLATFYQGKHIDANLDIHYGSSCEGQSSISIHGQYTHTDKDEEQLVN
1201 AAAGKPITGNLRYNGLHRMALQCNAGREHGIPFNYYCMKFLRHSSRLAKLTADVEWKNYK
1261 PLLNKLLPLHTKYHALRPEHGGFFGIIRSHFTGENGKLHLVSQVPWWDLKDKPHTDIVIT
1321 TEDGQHFKHWNVPTFSHMLEPRAFSSLGYSNIAEYAKQYRHRHCDLQKLSLRTFDGSLVQ
1381 LPETDCYKVVTRDCSPNKRFLVMARSTNNPSLTKALKVFIHTTKLEILPVTEDSGLIVRV
1441 DGNKVDVVPERPYSHTDHDVELFEVRTREKWFEVTSKSYGIYLTFNGNLLFIQTAPFYRG
1501 KLCGLCGDYNLDRNHELSGPDGHLYNSTLEFAKSYVVPSPDCHPPTH*
B
Secretion Signal
RXXR
Lipoprotein N-terminal Domain
DUF1943
vW Type D domain
▼= Predicted N-linked glycosylation site

Figure 2.

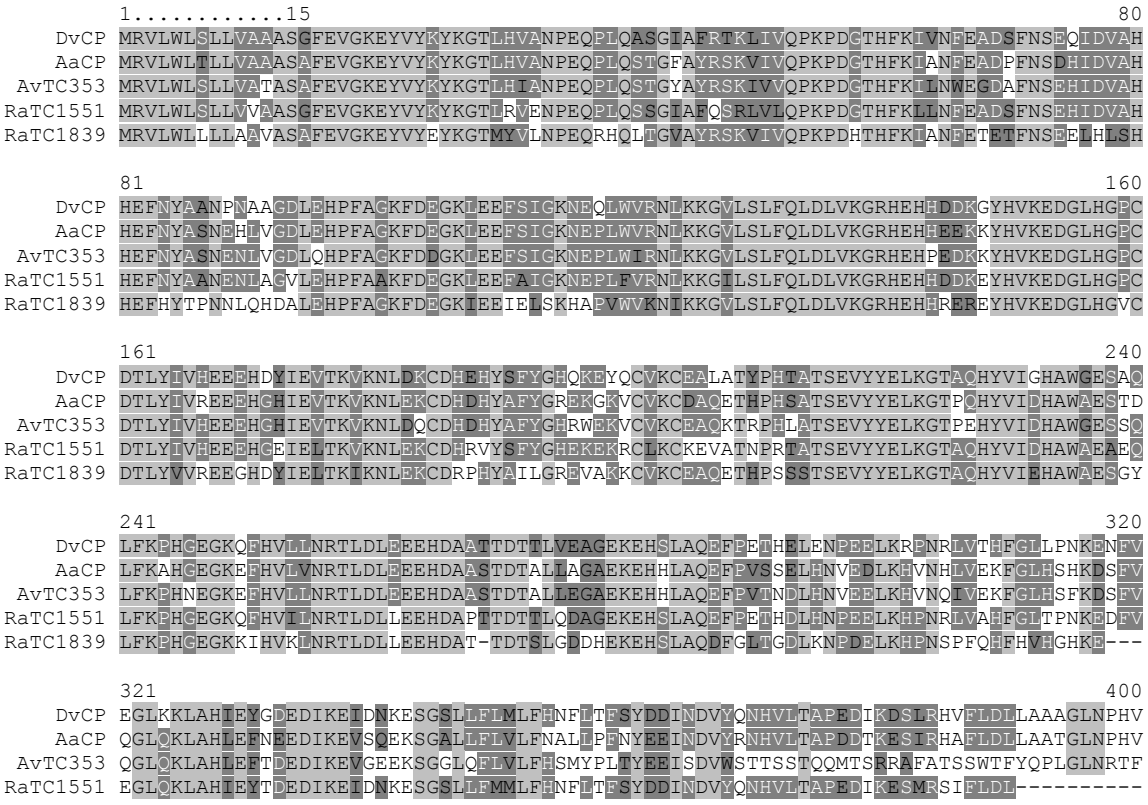

1.....15 80
DvCP MRVLWLSLLVAASGFEVGKEYVYKYKGTLHVANPEQPLQASGIAFRTKLIVQPKPDCTHFKIVNFEADSFNSEQIDVAH
AaCP MRVLWLTLLVAASAFEVGKEYVYKYKGTLHVANPEQPLQSTGFAYRSKVIVQPKPDCTHFKIANFEADPFNSDHIDVAH
AvTC353 MRVLWLSLLVATASAFEVGKEYVYKYKGTLHIANPEQPLQSTGYAYRSKIVVQPKPDCTHFKILNWEGDAFNSEHIDVAH
RaTC1551 MRVLWLSLLVVASGFEVGKEYVYKYKGTLRVENPEQPLQSSGIAFQSRLVLQPKPDCTHFKLLNFEADSFNSEHIDVAH
RaTC1839 MRVLWLLLLAAVASAFEVGKEYVYEYKGTMYVLNPEQRHQLTGVAYRSKVIVQPKPDHTHFKIANFETETFNSEELHLSH
81 160
DvCP HEFNYAANPNAAGDLEHPFACKFDEGKLEEFSICKNEQLWVRNLKKGVLSLFQLDLVKGRHEHHDDKGYHVKEDGLHGPC
AaCP HEFNYASNEHIVGDLEHPFACKFDEGKLEEFSICKNEPLWVRNLKKGVLSLFQLDLVKGRHEHHEEKKYHVKEDGLHGPC
AvTC353 HEFNYASNENIVGDLQHPFACKFDDGKLEEFSICKNEPLWIRNLKKGVLSLFQLDLVKGRHEHPEDKKYHVKEDGLHGPC
RaTC1551 HEFNYAANENIACVLEHPFAAKFDEGKLEEFAICKNEPLFVRNLKKGILSLFQLDLVKGRHEHHDDKEYHVKEDGLHGPC
RaTC1839 HEFHYTPNNLQHDALEHPFACKFDEGKIEEIELSKHAPVWVKNIKKGVLSLFQLDLVKGRHEHHRERFYHVKEDGLHGVC
161 240
DvCP DTLYIVHEEEHDYIEVTKVKNLDKCDHEHYSFYGHQKEYQCVKCEALATYPHTATSEVYYELKGTAQHYVIGHAWGESAQ
AaCP DTLYIVREEEHCHIEVTKVKNLEKCDHDHYAFYGREKGKVCVKCDAQETHPHSATSEVYYELKGTPQHYVIDHAWAESTD
AvTC353 DTLYIVHEEEHCHIEVTKVKNLDQCDHDHYAFYGHRWEKVCVKCEAQKTRPHLATSEVYYELKGTPEHYVIDHAWGESSQ
RaTC1551 DTLYIVHEEEHCEIELTKVKNLEKCDHRVYSFYGHEKEKRCLKCKEVATNPRTATSEVYYELKGTAQHYVIDHAWAEAEQ
RaTC1839 DTLYVVREEGHDYIELTKIKNLEKCDRPHYAILGRDVAKKCVKCEAQETHPSSSTSEVYYELKGTAQHYVIEHAWAESGY
241 320
DvCP LFKPHCEGKQGHVLLNRTLDLEEHDAATTDTTLVEAGEKEHSLAQEFPETHELENPEELKRPNRLVTHFGLLPNKENFV
AaCP LFKAHCEGKEGHVLVNRTLDLEEHDAASTDTALLAGAEKEHHLAQEFPVSSELHNVEDLKHVNHLVEKFGLHSHKDSFV
AvTC353 LFKPHNEGKEGHVLLNRTLDLEEHDAASTDTALLEGAEKEHHLAQEFPVTNDLHNVEELKHVNQIVEKFGLHSFKDSFV
RaTC1551 LFKPHCEGKQGHVILNRTLDLLEEHDAPTTDTTLQDAGEKEHSLAQEFPETHDLHNPEELKHPNRLVAHFGLTPNKEDFV
RaTC1839 LFKPHCEGKKIHVKLNRTLDLLEEHDAT-TDTSLGDDHEKEHSLAQDFGLTGDLKNPDELKHPNSPFQHFHVHGHKE---
321 400
DvCP EGLKKLAHIEYGDEDIKEIDNKESGSLLFLMLFHNFLTFSYDDINDVYQNHVLTAPEDIKDSLRHVFLDLLAAAGLNPHV
AaCP QGLQKLAHLEFNDEDIKEVSQEKSGALLFLVLFNALLPFNYEEINDVYRNHVLTAPDDTKESIRHAFLDLLAATGLNPHV
AvTC353 QGLQKLAHLEFTDEDIKEVGEEKSGGLQFLVLFHSMYPLTYEEISDVWSTTSSTQQMTSRRAFATSSWTFYQPLGLNRTF
RaTC1551 EGLQKLAHIEYTDEDIKEIDNKESGSLLFMMLFHNFLTFSYDDINDVYQNHVLTAPEDIKESMRSIFLDL----------

Figure 3.

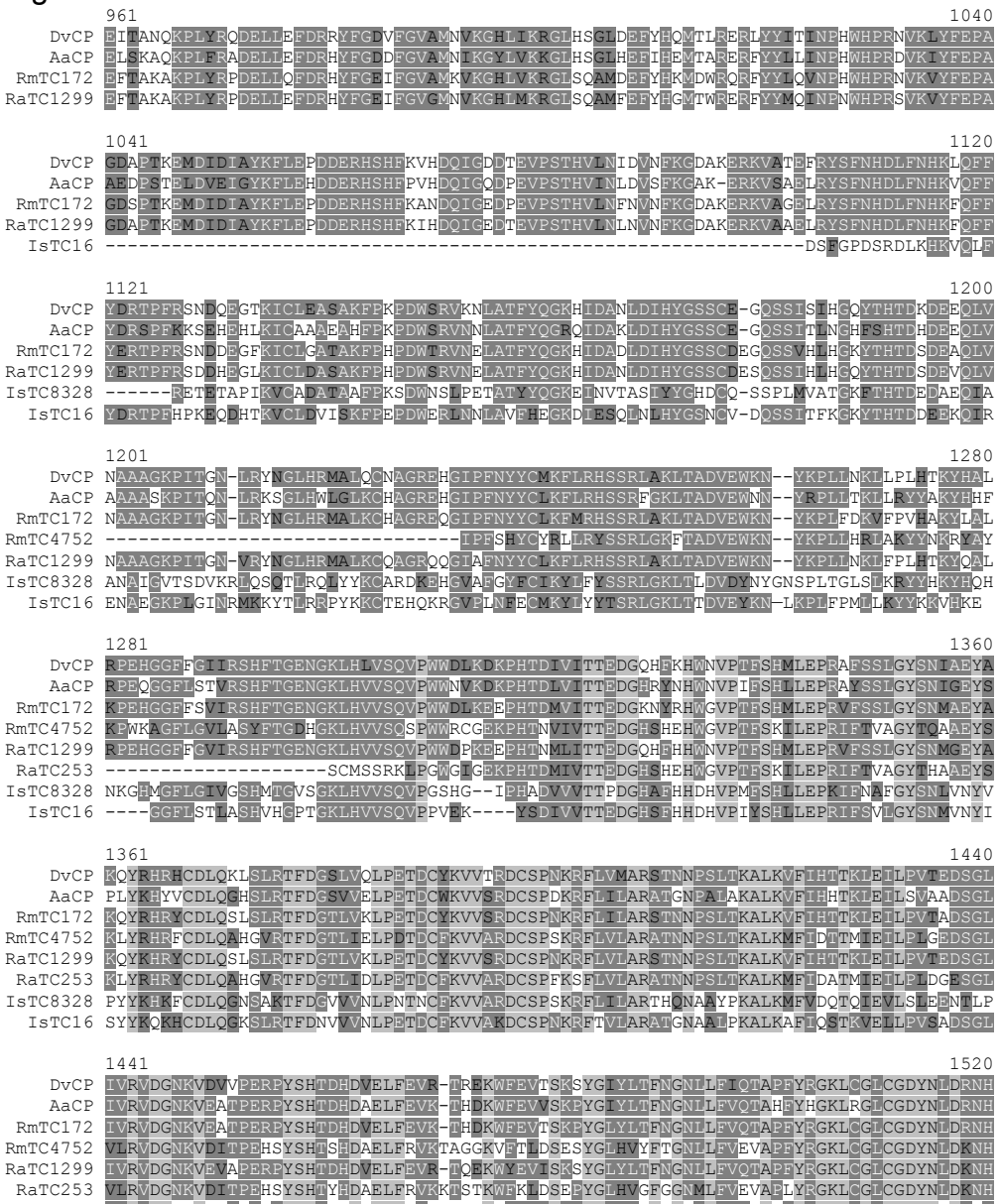

961 1040
DvCP EITANQKPLYRQDELLEFDRRYFGDVFGVAMNVKGHLIKRGLHSGLDEFYHQMTLRERLVYITINPHWHPRNVKLYFEPA
AaCP ELSKAQKPLFRADELLEFDRHYFGDDFGVAMNIKGYLVKKGLHSGLHEFIHEMTARERFYYLLINPHWHPRDVKIYFEPA
RmTC172 EFTAKAKPLYRPDELLQFDRHYFGEIFGVAMKVKGHLVKRGLSQAMDEFYHKMDWRQRFYYLQVNPHWHPRNVKVYFEPA
RaTC1299 EFTAKAKPLYRPDELLEFDRHYFGEIFGVGMNVKGHLMKRGLSQAMFEFYHGMTWRERFYYMQINPNWHPRSVKVYFEPA
1041 1120
DvCP GDAPTKEMDIDIAYKFLEPDDERHSHFKVHDQIGDDTEVPSTHVLNIDVNFKGDAKERKVATEFRYSFNHDLFNHKLQFF
AaCP AEDPSTELDVEIGYKFLEHDDERHSHFPVHDQICQDPEVPSTHVINLDVSFKGAK-ERKVSAELRYSFNHDLFNHKVQFF
RmTC172 GDSPTKEMDIDIAYKFLEPDDERHSHFKANDQIGDDPEVPSTHVLNFNVNFKGDAKERKVAGELRYSFNHDLFNHKFQFF
RaTC1299 GDAPTKEMDIDIAYKFLEPDDERHSHFKIHDQIGEDTEVPSTHVLNLNVNFKGDAKERKVAAELRYSFNHDLFNHKFQFF
IsTC16 ----------DSFGPDSRDLKHKVQLF
1121 1200
DvCP YDRTPFRSNDQEGTKICLEASAKFFKPDWSRVKNLATFYQGKHIDANLDIHYGSSCE-GQSSISIHGQYTHTDKDEEQLV
AaCP YDRSPFKKSEHEHLKICAAAEAHFFKPDWSRVNNLATFYQGRQIDAKLDIHYGSSCE-GQSSITLNGHFSHTDHDEEQLV
RmTC172 YERTPFRSNDDEGFKICLGATAKFFHPDWTRVNELATFYQGKHIDADLDIHYGSSCDEGQSSVHLHGKYTHTDSDEAQLV
RaTC1299 YERTPFRSDDHEGLKICLDASAKFFHPDWSRVNELATFYQGKHIDANLDIHYGSSCDESQSSIHLHGQYTHTDSDEVQLV
IsTC8328 ------RETETAPIKVCADATAAFFKSDWNSLPETATYYQGKELNVTASLYYGHDCQ-SSPLMVATGKFTHTDEDAEQIA
IsTC16 YDRTPFHPKEQDHTKVCLDVISKFFEPDWERLNNLAVFHEGKDLESQLNLHYGSNCV-DQSSITFKGKYTHTDDEEKQIR
1201 1280
DvCP NAAAGKPITGN-LRYNGLHRMALQCNAGREHGIPFNYYCMKFLRHSSRLAKLTADVEWKN--YKPLLNKLLPLHTKYHAL
AaCP AAAASKPITQN-LRKSGLHWLGLKCHAGREHGIPFNYYCLKFLRHSSRFGKLTADVEWNN--YRPLLTKLLRYYAKYHHF
RmTC172 NAAAGKPITGN-LRYNGLHRMALKCHAGREQGIPFNYYCLKFMRHSSRLAKLTADVEWKN--YKPLFDKVFPVHAKYLAL
RmTC4752 ----------IPFSHYQYRLLRYSSRLGKFTADVEWKN--YKPLLHRLAKYYNKRYAY
RaTC1299 NAAAGKPITGN-VRYNGLHRMALKCQAGRQQGIAFNYYCLKFLRHSSRLAKLTADVEWKN--YKPLLNKLFPLHTKYQAL
IsTC8328 ANAICVTSDVKRLQSQTLRQLYYKCARDKFHGVAFGYFCIKYLFYSSRLGKLTLDVDYNYGNSPLTGLSLKRYYHKYHQH
IsTC16 ENAEGKPLGINRMKKYTLRFPYKKCTEHQKRGVPLNFECMKYLYYTSRLGKLTTDVEYKN-LKPLFPMLLKYYKKVHKE
1281 1360
DvCP RPEHGGFFGIIRSHFTGENGKLHLVSQVPWWDLKDKPHTDIVITTEDGQHFKHWNVPTFSHMLEPRAFSSLGYSNIAEYA
AaCP RPEQGGFLSTVRSHFTGENGKLHVVSQVPWWNVKDKPHTDLVITTEDGHRYNHWNVPIFSHLLEPRAYSSLGYSNIGEYS
RmTC172 KPEHGGFFSVIRSHFTGENGKLHVVSQVPWWDLKEEPHTDMVITTEDGKNYRHWGVPTFSHMLEPRVFSSLGYSNMAEYA
RmTC4752 KPWKAGFLGVLASYFTGDHGKLHVVSQSPWWRCGEKPHTNVIVTTEDGHSHEHWGVPTFSKILEPRIFTVAGYTQAAEYS
RaTC1299 RPEHGGFFGVIRSHFTGENGKLHVVSQVPWWDPKEEPHTNMLITTEDGQHFHHWNVPTFSHMLEPRVFSSLGYSNMGEYA
RaTC253 ----------SCMSSRKLPGWGIGEKPHTDMIVTTEDGHSHEHWGVPTFSKILEPRIFTVAGYTHAAEYS
IsTC8328 NKGHMGFLGIVGSHMTGVSGKLHVVSQVPGSHG--IPHADVVVTTPDGHAFHHDHVPMFSHLLEPKIFNAFGYSNLVNYV
IsTC16 ----GGFLSTIASHVHGPTGKLHVVSQVPPVEK----YSDIVVTTEDGHSFHHDHVPIYSHLLEPRIFSVLGYSNMVNYI
1361 1440
DvCP KQYRHRHCDLQKLSLRTFDGSLVQLPETDCYKVVTRDCSPNKRFLVMARSTNNPSLTKALKVFIHTTKLEILPVTEDSGL
AaCP PLYKHYVCDLQGHSLRTFDGSVVELPETDCWKVVSRDCSPDKRFLILARATGNPALAKALKVFIHHTKLEILSVAADSGL
RmTC172 KQYRHRYCDLQSLSLRTFDGTLVKLPETDCYKVVSRDCSPNKRFLILARSTNNPSLTKALKVFIHTTKLEILPVTADSGL
RmTC4752 KLYRHRFCDLQAHGVRTFDGTLIELPDTDCFKVVARDCSPSKRFLVLARATNNPSLTKALKMFIDTTMIEILPLGEDSGL
RaTC1299 KQYKHRYCDLQSLSLRTFDGTLVKLPETDCYKVVSRDCSPNKRFLVLARSTNNPSLTKALKVFIHTTKLEILPVTEDSGL
RaTC253 KLYRHRYCDLQAHGVRTFDGTLIDLPETDCFKVVARDCSPFKSFLVLARATNNPSLTKALKMFIDATMIEILPLDGESGL
IsTC8328 PYYKHKFCDLQGNSAKTFDGVVVNLPNTNCFKVVARDCSPSKRFLILARTHQNAAYPKALKMFVDQTQIEVLSLEENTLP
IsTC16 SYYKQKHCDLQGKSLRTFDNVVVNLPETDCFKVVAKDCSPNKRFTVLARATGNAALPKALKAFIQSTKVELLPVSADSGL
1441 1520
DvCP IVRVDGNKVDVVPERPYSHTDHDVELFEVR-TREKWFEVTSKSYGIYLTFNGNLLFIQTAPFYRGKLCGLCGDYNLDRNH
AaCP IVRVDGNKVEATPERPYSHTDHDAELFEVK-THDKWFEVVSKPYGIYLTFNGNLLFVQTAHFYHGKLRGLCGDYNLDRNH
RmTC172 IVRVDGNKVEATPERPYSHTDHDVELFEVK-THDKWFEVTSKPYGLYLTFNGNLLFVQTAPFYRGKLCGLCGDYNLDRNH
RmTC4752 VLRVDGNKVDITPEHSYSHTSHDAELFRVKTAGGKVFTLDSESYGLHVYFTGNLLFVEVAPFYRGKLCGLCGDYNLDKNH
RaTC1299 IVRVDGNKVEVAPERPYSHTDHDVELFEVR-TQEKWYEVISKSYGLYLTFNGNLLFVQTAPFYRGKLCGLCGDYNLDKNH
RaTC253 VLRVDGNKVDITPEHSYSHTYHDAELFRVKKTSTKWFKLDSEPYGLHVGFGGNMLFVEVAPLYRGKLCGLCGDYNLDKNH
IsTC8328 IVRVDGSQVAPSLESFYSHSVADAELFTISLMSNKNFEIVSKPYGIYVLFDGRILYVQVAHFYRGKLCGICGDYNYDRHH
IsTC16 VLRIDGNKVEVTPGVFYSHQAHDVELFTVT-QHNKYFEVMSQPYGVFIGFDGNVLFVQTANFYRGKLCGLCGDYNYDRST
1521 1552
DvCP ELSGPDGHLYNSTLEFAKSYVVPSPDCHPETH
AaCP ELSGPDGRHYNNSLEFAKSYVVPSPDCHAPAH
RmTC172 ELSGPDGHLYNNTLEFAKSYVVPSPECQAPAH
RmTC4752 ELSGPDGHQYNNTLEFAKSYVVPSPDCHAPAH
RaTC1299 DLSGPDGHLYNNTLEFAKSYVVPSPDCHAPAH
RaTC253 ELSGPDGHLYNNTLEFAKSYVVPSPDCHAPAH
IsTC8328 ELVGENLHLYNDSLQFAASYVVHSSDCTSEE-
IsTC16 SSWPQPSTTTNEHPGVRQGATFGSA-------

Figure 4.

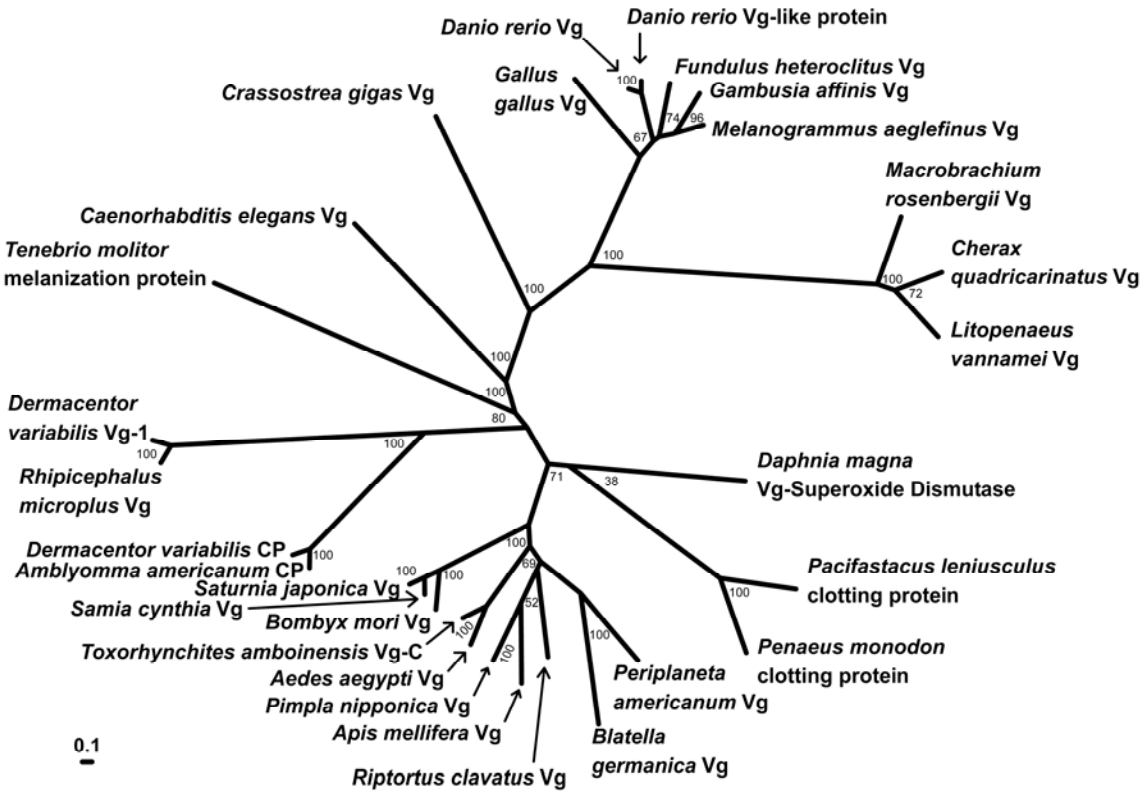

Danio rerio Vg
Danio rerio Vg-like protein
Fundulus heteroclitus Vg
Gambusia affinis Vg
Gallus gallus Vg
Crassostrea gigas Vg
Melanogrammus aeglefinus Vg
Macrobrachium rosenbergii Vg
Caenorhabditis elegans Vg
Tenebrio molitor melanization protein
Cherax quadricarinatus Vg
Litopenaeus vannamei Vg
Dermacentor variabilis Vg-1
Rhipicephalus microplus Vg
Daphnia magna Vg-Superoxide Dismutase
Dermacentor variabilis CP
Amblyomma americanum CP
Pacifastacus leniusculus clotting protein
Saturnia japonica Vg
Samia cynthia Vg
Bombyx mori Vg
Toxorhynchites amboinensis Vg-C
Penaeus monodon clotting protein
Aedes aegypti Vg
Periplaneta americanum Vg
Pimpla nipponica Vg
Apis mellifera Vg
Blatella germanica Vg
Riptortus clavatus Vg
0.1

Figure 5.

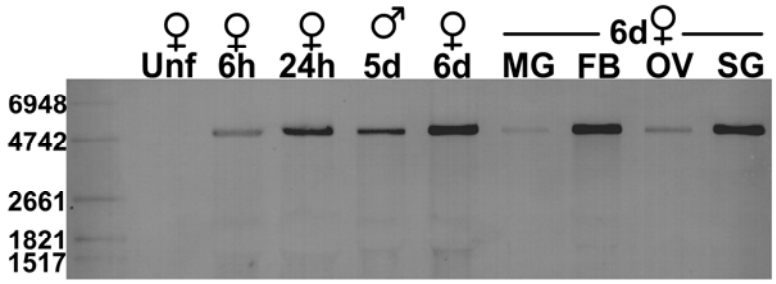
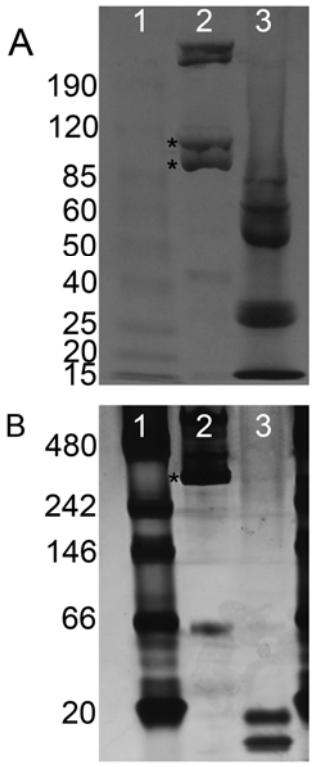


Figure 6.



Gene expression of neuropeptides identified by pyrosequencing of the American Dog tick synganglion transcriptome during blood-feeding and reproduction

Abstract

Ticks are important vectors of numerous pathogens that impact human and animal health. The tick central nervous system represents an understudied area in tick biology and no tick synganglion-specific transcriptome has been described to date. Here we characterize fourteen putative neuropeptides (allatostatin, insulin-like peptide, ion-transport peptide, sulfakinin, bursicon alpha/beta, eclosion hormone, glycoprotein hormone alpha/beta, corazonin, four orcokinins) and five neuropeptide receptors (gonadotropin receptor, leucokinin-like receptor, sulfakinin receptor, calcitonin receptor, pyrokinin receptor) from the synganglion of female American dog ticks, *Dermacentor variabilis*. Their homology to the same neuropeptides in other taxa will be discussed. Many of these neuropeptides such as an allatostatin, insulin-like peptide, eclosion hormone, bursicon alpha and beta and glycoprotein hormone alpha and beta have not been previously described in the Chelicerata. An insulin receptor substrate protein was also found indicating that at insulin signaling network is present in ticks. A putative type-2 proprotein processing convertase was also sequenced that may be involved in cleavage at monobasic and dibasic endoproteolytic cleavage sites in prohormone peptides. Quantitative real-time PCR was used to monitor developmental expression of these genes during adult female reproduction. Their physiological role during adult tick blood feeding and reproduction will be discussed.

Introduction

Ticks are obligate ectoparasites that vector the greatest variety of pathogens of any blood-feeding arthropod (Sonenshine, 1993). Development of precisely targeted tick control strategies could be ameliorated by a better understanding of the basic physiology of these organisms. The processes that control tick feeding and reproduction represent obvious targets for developing such a strategy. However, only recently have the major yolk proteins been fully sequenced, and only a few species have been studied (Thompson et al. 2007, Mitchell et al. 2006). Furthermore, other recent studies on vitellogenesis in ticks indicate that a consensus model of tick reproduction will require studies from multiple taxa in order to understand key regulatory differences (Seixas et al. 2008).

Blood-feeding and mating are necessary requirements for the development of viable eggs in ticks. However, the timing of blood-feeding with that of mating differs among the Prostriata, such as *Ixodes scapularis*, the blacklegged tick, and the Metastriata, such as *Dermacentor variabilis*, the American dog. In metastriate ticks, blood feeding is initiated with attachment to the host and once mated, female ticks engorge to repletion. Evidence suggests that the synganglion produces a prothoracicotropic hormone (PTTH)-like hormone that triggers 20-hydroxyecdysone (20-HE) synthesis by the epidermis. Recent studies have shown that 20-HE initiates expression of the major yolk protein, vitellogenin (Vg), by the fat body and to a lesser extent the midgut (Thompson et al. 2006, 2007). Deposition of vitellogenin into the oocytes via the vitellogenin receptor results in vitellin formation and subsequently oviposition (Mitchell et al. 2007). Seixas and colleagues (2008) demonstrated

that in *Amblyomma hebraeum* 20-HE triggers Vg expression but not deposition into the oocytes suggesting that a "vitellogenin factor" in the hemolymph is required. Regardless, in both species blood feeding initiates the cascade of events that likely results in the release of a PTTH-like substance from the synganglion that triggers 20-HE synthesis by the epidermis although this hormone has not been identified.

Published transcriptome data from the tick synganglion is lacking due in part to the difficulty of extracting sufficient quantities of tissue. Previous to the recent bioinformatics based work of Christie (2008), only a few hormones in the Chelicerata were known, and only at the protein level obtained from direct protein sequencing. In the current study we used pyrosequencing to characterize the transcribed genes produced from pooled adult tick synganglion samples taken from unfed, partfed and replete females of the American dog tick with the goal of identifying neuropeptides and associated receptors.

Materials and methods

Ticks

American dog ticks, *Dermacentor variabilis*, were reared as previously described (Sonenshine, 1993) and originated from specimens collected near Richmond, Virginia, USA. Adult ticks were confined within plastic capsules attached to New Zealand white rabbits, *Oryctolagus cuniculus*, and allowed to feed as required for experiments described below. Larvae and nymphs were fed on Norway rats, *Rattus norvegicus*. Rearing conditions were 26 ± 1°C, 92 ± 6% relative humidity and 14:10 (L:D).

454 library preparation

Approximately 50 synganglia for each developmental condition, including lateral secretory organs and pedal nerves were dissected from female *D. variabilis* that were either unfed, part-fed virgin (attached to the host for 4-5 d), part-fed virgin forcibly detached from the rabbit host and held in culture for 4-5 d, part-fed mated (allowed to mate for ≤1 d) or replete (1 d post drop off from the host). Tissues from each feeding stage were separately homogenized in TRI Reagent (Sigma-Aldrich, St. Louis, MO) according to the manufacturer's recommendations. RNA pellets were rehydrated in 100 µM aurintricarboxylic acid to prevent degradation (Hallick et al., 1977). Approximately 3 µg of total RNA from each group were pooled, and mRNA was isolated using an Oligotex mRNA isolation kit (Qiagen, Valencia, CA) according to the manufacturer's recommendations. Purified mRNA was ethanol precipitated and rehydrated in 2 µl and combined with 10 picomoles of modified 3' reverse transcription primer (5'-ATTCTAGAGACCGAGGCGGCCGACATGT$_{(4)}$GT$_{(9)}$CT$_{(10)}$VN-3') (Beldade et al. 2006) and 10 picomoles SMART IV oligo (5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3') (Zhu et al. 2001). The resulting 4 µl were incubated at 72°C for 2 min and then combined with the following reagents on ice: 1 µl RNase Out (40 U/µl), 2 µl 5X first strand buffer, 1 µl 20 mM DTT, 1 µl dNTP mix (10 mM each) and 1 µl Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA). The reaction was incubated at 42°C for 90 min then diluted to 30 µl with TE buffer (10 mM Tris HCL pH 7.5, 1 mM EDTA) and stored at -20°C until further use. To synthesize

second strand cDNA, 5 µl of first-strand cDNA was mixed with 10 picomoles of modified 3' PCR primer (5'-ATTCTAGAGGCCGAGGCGGCCGACATGT$_{(4)}$GTCT$_{(4)}$GTTCTGT$_{(3)}$CT$_{(4)}$VN-3') (Beldade et al. 2006), 10 picomoles of 5' PCR primer (5'-AAGCAGTGGTATCAACGCAGAGT-3') (Zhu et al. 2001), 5 µl 10X reaction buffer, 1 µl dNTP mix, 2 µl $MgSO_4$, 0.4 µl Platinum HiFi Taq Polymerase and 34.6 µl H_2O (Invitrogen). Thermal cycling conditions were 94°C for 2 min followed by 20 cycles of 94°C for 20 sec, 65°C for 20 sec and 68°C for 6 min. The first PCR reaction was conducted, and 5 µl aliquots from cycles 18, 22 and 25 were analyzed on a 1% agarose gel to optimize the number of cycles. An additional 5 reactions were conducted to produce sufficient quantities of cDNA for 454 library preparation. The contents were combined, and the cDNA was purified by using a PCR purification kit (Qiagen) according to the manufacturer's recommendations. The cDNA library was prepared with appropriate kits (Roche, Indianapolis, IN; Qiagen) for pyrosequencing on the GS-FLX sequencer (Roche) according to the manufacturer's recommendations which were described previously (Margulies et al. 2005). The only deviation from the protocol was prior to titration sequencing; following emulsification PCR, DNA-positive beads were enriched in order to increase the number of reads collected during titration.

Quantitative real-time PCR

Four groups of 20 synganglia each were dissected and pooled separately from unfed, part-fed and replete female *D. variabilis* and immediately placed in extraction buffer (Molecular Devices, Sunnyvale, CA) on dry ice. DNase treated (Qiagen) total RNA was isolated with the PicoPure kit (Molecular Devices) according to the manufacturer's protocol. One microgram of total RNA from each sample was linearly amplified using the MessageAmp II aRNA amplification kit (Ambion). Five hundred nanograms of aRNA were reverse transcribed with 200 ng of a random pentadecamer and superscript II reverse transcriptase. cDNA was diluted 1:2 in H_2O, 2 µl were used per reaction for qRT-PCR with SYBR green master mix (Applied Biosystems), and reactions were carried out on a ABI Prism 7900 sequence detector (Applied Biosystems). Amplified products were normalized to GAPDH (EU999993) and analyzed using the $2^{-\Delta\,\Delta C}{}_T$ method (Livak et al. 2001). Primer sequences are available upon request. Comparison of unfed, part-fed and replete quantitative real-time polymerase chain reaction (qRT-PCR) data (ΔCt) were analyzed initially using an analysis of variance (PROC GLM) using SAS ver. 9 (SAS Institute, Cary, NC). Pairwise comparisons of expression data using least squares means were assessed with a Tukey's test (PROC GLM).

Bioinformatics

Removal of primer sequence contamination and assembly of GS-FLX sequencing reads were carried out with GS Assembler ver. 1.1.02.15 (Roche) using default parameters as

follows; seed step: 12, seed length: 16, seed count: 1, minimum overlap length: 40, minimum overlap identity: 90%, alignment identity score: 2, alignment difference score: -3. Assembled contiguous sequences, herein referred to as contigs, were initially identified using the Tera-BLASTX algorithm with DeCypher (TimeLogic) against Genbank nr and est databases (downloaded June 2008). BLAST searches (Altschul et al. 1997) against the *Ixodes scapularis* genomic contigs (ver. IscaW1) and predicted transcripts (ver. IscaW1.05.1) were performed at the Vectorbase website (www.vectorbase.org). Sequence alignments were performed with ClustalW (www.ebi.ac.uk/clustalw). Heat map construction of qRT-PCR fold data was carried out with the online tool Matrix2png (http://www.bioinformatics.ubc.ca/pavlidis/lab/cgi-bin/matrix2png.cgi). Phylogenetic trees were constructed with Mega4 (Tamura et al. 2007). Secretion signal prediction was carried out at the SignalP 3.0 server website (http://www.cbs.dtu.dk/services/SignalP), and searches were made using both neural networks and hidden markov models (Bendtsen et al. 2004).

Results

Allatostatin type-A

An allatostatin was putatively identified from *D. variabilis* as a type-A allatostatin (ACC99603). The 5'-truncated sequence is 430 bp and codes for a 108-residue peptide that was used to locate a similar gene from the recent draft sequence of the blacklegged tick, *I. scapularis* (ABJB010344865), an EST from the Southern cattle tick, *Rhipicephalus microplus* (CK192037) and the Chilean recluse spider, *Loxosceles laeta* (EY189417)(Figure

1). The *I. scapularis* allatostatin is full length and is coded for by a single exon. The translation of this sequence results in a 169 residue protein with a 22 amino acid secretion signal. The 3'-partial *R. microplus* EST translates to a 157 residue protein lacking the carboxy terminus and contains a 36-residue secretion signal. The *L. laeta* sequence also lacks the 5' region of the transcript and encodes for a 111 amino acid protein.

All four sequences contain dibasic endoproteolytic cleavage sites that flank the predicted mature peptides that contain a carboxy-terminal glycine residue that would presumably undergo amidation. The complete *I. scapularis* sequence would yield four isoforms while the *L. laeta* sequence would produce a fifth peptide due to a Lys/Arg cleavage site not present in the tick sequences.

Gene expression of allatostatin in the synganglia of unfed, part-fed and replete female ticks was measured by qRT-PCR (Figure 2). The mRNA levels were 3.22- and 2.43-fold greater in unfed ticks than in part-fed (Tukey's test, $P<0.0003$) and replete ticks ($P<0.0018$), respectively. Expression of allatostatin in part-fed ticks was not significantly different when compared to replete ticks ($P=0.31$).

Insulin-like peptide

A 171 bp contig from the *D. variabilis* synganglion library was identified as an insulin-like peptide (ACC99597) by BLASTx (Altschul et al. 1997) against the nr Genbank database. Translation of the 5'-partial sequence returns a 22-residue peptide that represents the A-chain of an insulin-like peptide (Figure 3A). No acarine EST data are available that

are homologous with *D. varabilis* insulin-like peptide (DvILP); however a highly homologous predicted transcript that was 90.9% identical in the aligned region was located in the *I. scapularis* genome (ABJB010768667). A protein query against the nr Genbank database returns results from insects and fish of several species. An alignment of the predicted A-chains of these sequences with that of DvILP and *I. scapularis* ILP (IsILP) demonstrates that these peptides are highly conserved across diverse taxa (Figure 3A). IsILP contains a single 1078 bp intron and is located in the region that codes for the C-chain. The predicted protein contains a 25-residue secretion signal and Lys/Arg cleavage sites that would result in mature B-, C- and A-chains.

Expression of DvILP was significantly higher in unfed female ticks than part-fed (Tukey's test, $P<0.03$) and replete ticks ($P<0.001$) (Figure 2). There was no significant difference between part-fed and replete ticks ($P=0.084$).

Gonadotropin releasing hormone receptor/Cardioacceleratory peptide receptor

A 68-residue peptide translated from a 207-bp contig was putatively identified as a gonadotropin releasing hormone receptor (ACC99628) by BLASTx with a top match to the zebrafish, *Danio rerio* (XP_697400; BLAST score, 53.1; E-value, 7e-6). The majority of the other top 100 matches by BLASTx were also gonadotropin-releasing hormone receptors. A search of the Pfam database for conserved motifs yields an insignificant match to the 7 transmembrane receptor rhodopsin family (7tm_1; Pfam PF00001) with an E-value of 0.014. A search against available acarine EST data indicated a match to *I. scapularis* EST

EW846470. A BLASTp search of the translation of EW846470 indicates that a 7tm_1 motif is present (E-value, 1e-19), and the top non-predicted match in the nr Genbank database is a gonadotropin-releasing hormone receptor 2 from the sea lamprey, *Petromyzon marinus* (AB077118; BLAST score, 108; E-value, 3e-22). Using the *D. variabilis* translation as a query against the *I. scapularis* genome returns a 297-residue predicted peptide found within the scaffold DS713552. This predicted peptide when used as a BLASTp query against the nr database appears to be a cardioacceleratory peptide receptor due to a match to the same peptide from *Aedes aegypti* (XP_001659389; Score, 130; E-value, 1e-28). An alignment of the *D. variabilis* and *I. scapularis* sequences suggests that the *D. variabilis* gene may be a cardioacceleratory peptide receptor and not a gonadotropin releasing hormone receptor (Figure 4). Expression of the *D. variabilis* receptor was 3.2- and 3.1-fold higher in unfed female synganglion over that of part-fed (Tukey's test, $P<0.03$) and repletes ($P<0.03$), and did not differ significantly between part-feds and repletes ($P=0.99$) (Figure 2).

Ion-transport peptide

The complete sequence for an ion-transport like peptide (ITP) was present in our library (EU620224). The 1193 bp contig codes for a 114-residue preproprotein. The 5' untranslated region (UTR) is 180 bp and the 3' UTR is 671 bp. The resulting protein contains a 33 amino acid secretion signal and contains a dibasic cleavage site at Lys^{38}/Arg^{39} that after processing would result in a 75-residue mature peptide. Christie (2008) reported a similar sequence from *I. scapularis* EST EW937910. The coding sequence for the *I.*

scapularis gene is located on scaffold DS934076 in Genbank and consists of two exons. Alignment of *D. variabilis* and *I. scapularis* ITP reveals that they are 68.4% identical. The *I. scapularis* ITP contains a dibasic cleavage site in the same location as in *D. variabilis* ITP after secretion signal removal and one near the carboxy-terminus (Figure 5A) which is not found in the *D. variabilis* translation. This results in an additional peptide after proprotein processing in *I. scapularis* than in *D. variabilis*. The implication of this in terms of the function of these proteins is currently unknown.

The *D. variabilis* and *I. scapularis* proteins both contain a crustacean hyperglycemic hormone/molt-inhibiting hormone/gonad-inhibiting hormone (CHH/MIH/GIH) motif (Pfam PF011147). A BLASTp search against the Genbank nr database returns multiple hits to ion-transport peptide and crustacean hyperglycemic hormones with the majority of the ITPs resulting in more significant expect values (E-values, ≤3e-11) than with CHHs (E-values, ≥5e-11). A phylogenetic analysis with the top 100 BLASTp results suggests that these tick peptides are more similar to ITPs than CHHs (Figure 5B).

Expression of the DvITP gene was highest in unfed female tick synganglia compared to part-fed and replete ticks (Figure 2). Levels in part-fed and replete female ticks were significantly less than unfed ticks by 3.85- and 2.85-fold (Tukey's test, $P<0.008$ and 0.03), respectively. Part-fed and replete ticks were not significantly different ($P=0.65$).

Sulfakinin and sulfakinin receptor

Both a sulfakinin and a sulfakinin-like receptor were sequenced. The putative sulfakinin is a 648 bp transcript that encodes a full length 115-residue precursor protein (ACC99604). The cDNA contains a 46 bp 5' UTR and a 257 bp 3' UTR. The preproprotein contains a 21-residue secretion signal and two dibasic cleavage sites at Lys^{90} /Arg^{91} and the C-terminally located Arg^{114}/Lys^{115}. A monobasic cleavage site at Arg^{133} is also predicted. These features are in agreement with a putative *I. scapularis* sulfakinin reported by Christie (2008). An alignment of *D. variabilis* and *I. scapularis* sulfakinins (Figure 6A) indicates that the mature peptides flanked by endoproteolytic cleavage sites differ only by a single residue that is an aspartic acid in *D. variabilis* and a glutamic acid in *I. scapularis*. The upstream region that gives rise to the largest of the three mature peptides is 48.5% identical.

The two mature peptides QEDDYGHMRFG and SDDYGHMRFG are predicted to undergo extensive posttranslational modification (Christie 2008). The glutamine residue in the former may cyclize and the C-terminal glycine residues would likely undergo amidation. Analysis with the program Sulfinator predicts that the tyrosine residues in each peptide would be sulfated (E-values, 1.5; threshold=55).

A partial sequence for a *D. variabilis* sulfakinin receptor (ACC99631) was sequenced. The 228 bp fragment translates to a 75-residue protein that lacks both the start and stop codons. BLASTp returns the best match to the perisulfakinin receptor from the American cockroach, *Periplaneta americana* (AAX56942, BLAST score 68.9; E-value, 1e-10). A 7tm_1 motif is also predicted within this region (E-value, 4e-4). Using the *D.*

variabilis sequence as a query, an *I. scapularis* EST (EW798836) was located which translates to a 173-residue peptide lacking the complete N-terminus (Figure 6B). The translation is predicted to contain a 7tm_1 motif (E-value, 1e-9) like that found in the *D. variabilis* receptor.

Gene expression of profiles of the sulfakinin and sulfakinin receptor were similar in that the highest levels were found in unfed females (Figure 2). Sulfakinin was 3.22- and 5.55-fold higher in unfed ticks than in part-fed and repletes (Tukey's test, $P<0.003$ and $P<0.002$, respectively). The levels between part-feds and repletes were not significantly different ($P=0.12$). A similar result was found for the sulfakinin receptor with levels in unfed females that were 4.35- and 3.22-fold greater than in part-fed and replete ticks, respectively ($P<0.002$ and $P<0.007$, respectively). Levels in part-fed and repletes were not significantly different ($P=62$).

Leucokinin-like and calcitonin-like receptors

The incomplete sequence of a leucokinin-like receptor (ACC99630) in *D. variabilis* was identified by comparison to the same gene reported in *Rhipicephalus microplus* (AAF72891). The 227 bp *D. variabilis* transcript contained no start or stop codon and is translatable to a 76-residue incomplete peptide that contains a partial 7tm_1 motif. The *D. variabilis* and *R. microplus* sequences are 100% identical with the exception of Ile^{64} in *R. microplus* being a Leu residue in *D. variabilis*. Synganglion-specific gene expression (Figure 2) was 16.6- and 7.7-fold higher in unfed females compared to part-fed and replete

females (Tukey's test, $P<0.003$ and $P<0.02$, respectively). Gene expression did not differ significantly between part-fed and replete ticks ($P=0.41$).

A 377 bp, 5' and 3' partial *D. variabilis* calcitonin-like receptor transcript (ACC99624) was identified that encodes a 114 amino acid peptide. A search against the Pfam database indicates the presence of a 7-transmembrane family 2 motif (7tm_2) that is N-terminally truncated and extends to Val^{61}. A search against the Genbank nr database by BLASTp returns the top non-predicted result as calcitonin-receptor from the fish, *Gasterosteus aculeatus* (FAA00374; BLAST score, 73.9; E-value, 4e-12). The majority of the top 100 returned results were also calcitonin receptors from vertebrate and non-vertebrate taxa. No acarine homolog was located in any EST database although one was located in the *I. scapularis* genome (DS922272). The partial peptide sequence returns only calcitonin receptors when used as a query against the Genbank nr database by BLASTp. Furthermore, the predicted peptide also contains a 7tm_2 motif as well as a hormone receptor family motif (Pfam, PF02793). When aligned, the *D. variabilis* and *I. scapularis* sequences are 58% identical in the 68-residue region of overlap (data not shown).

Expression of the *D. variabilis* calcitonin-like receptor gene increased with attachment to the host and then appeared to decrease with mating (Figure 2). Transcript levels in female synganglia were 2.28-fold higher in part-fed than unfed females (Tukey's test, $P<0.03$). No significant difference was found between unfed and replete ticks ($P=0.43$) and part-fed and replete ticks ($P=0.22$).

Eclosion hormone

Eclosion hormone was putatively identified from a 5' partial 224 bp transcript that encodes a 66-residue peptide. Comparison against the Pfam database indicates the presence of an eclosion hormone motif (Pfam PF04736) that extends from residues 12 to 64. When used as a query against the Genbank nr database, all results matched with eclosion hormones from insects with the best match to a putative *Aedes aegypti* eclosion hormone (BLAST score, 57.4; E-value, 3e-7). No available EST data matched with *D. variabilis* eclosion hormone; however, a single exon from the *I. scapularis* genome encodes a putative 5' partial eclosion hormone (ABJB010327232). The upstream region of the *I. scapularis* gene could not be located. Translation of the *I. scapularis* putative eclosion hormone indicates that this peptide also contains an eclosion hormone motif. The expression in *D. variabilis* synganglion did not change significantly in unfed, part-fed or replete females (ANOVA, P=0.14)(Figure 2). However, the data appeared to suggest based on means alone, a decrease in the transcript levels in part-fed ticks. This may suggest that comparing unfed, part-fed and replete ticks does not offer the resolution needed to account for rapid changes in gene expression at precise time points during attachment to the host and mating.

Bursicon

Homologs to the alpha and beta subunits of the insect tanning hormone bursicon were identified. *D. variabilis* bursicon alpha (bur-α; ACC99596) is encoded by a 223 bp, 5' partial sequence that translates to a 69-residue peptide; bursicon beta (bur-β; ACC99598) is

encoded by a 515 bp, 3' partial sequence with a 236 5' UTR. The translation of bur-β results in a 93-residue peptide with a 21-residue secretion signal. Full length bur-α and bur-β genes are located in trans on the *I. scapularis* genome (DS725348), and each is the product of two exons. Using *I. scapularis* bur-α (Isbur-α) and bur-β (Isbur-β) as BLASTp queries against the nr database returns the top match to bursicon the crustacean, *Daphnia arenata* in both cases (Isbur-α: BLAST score, 152; E-value, 9e-37; Isbur-β: BLAST score, 119; E-value, 8e-26).

Surprisingly, the expression profiles of the bursicon subunits in *D. variabilis* were not identical (Figure 2). Dvbur-α gene expression was stable in unfed, part-fed and replete ticks (ANOVA, P=0.58) while Dvbur-β gene expression increased 4.22-fold from unfed to part-fed ticks (P<0.0004) and 12.8-fold from unfed to replete ticks (P<0.0001). Expression in replete ticks was 3.03-fold greater than part-fed ticks (P<0.002).

Glycoprotein hormone alpha/beta

Glycoprotein hormone-α (gh-α) and -β (gh-β) were sequenced in the *D. variabilis* synganglion library. Dvgh-α (ACC99601) is a 617 bp, 3' partial transcript that encodes a 116-residue peptide that contains 21-residue secretion signal. BLASTp results indicate the top match is gh-α2 from *Culex quinquefasciatus* (BLAST score, 133; E-value, 3e-30). A TBLASTn search against the EST database returns an *I. scapularis* EST (EL516156) sequenced from synganglia that is 96% identical (BLAST score, 153; E-value, 1e-37) and a

second *I. scapularis* EST (EW905646) that is 80% identical (BLAST score, 145; E-value, 2e-37).

Dvgh-β (ACC99600) is a 198 bp, 5' partial sequence that encodes a 50-residue peptide. Using this peptide as a query with a BLASTp search shows the best match is gh-β from *Anopheles gambiae* (XP_555160; BLAST score 38.5; E-value, 0.15). All other matches were also gh-βs or unnamed records. No EST record matched Dvgh-β, but the gene could be located in the *I. scapularis* genome (DS860962). The *I. scapularis* gh-β (Isgh-β) sequence is 78% identical to Dvgh-β in overlapping regions. Furthermore Isgh-β when translated is predicted by SignalP-HMM to contain a 16-residue secretion signal, while SignalP-NN does not agree and does not predict a secretion signal.

Gene expression of Dvgh-α was 4.34- and 4.55-fold greater in part-fed and replete tick synganglia over that expressed in unfed female ticks (Tukey's test, $P<0.0001$ and $P<0.0001$, respectively (Figure 2). Expression levels between part-fed and replete ticks was not significantly different ($P=0.76$). Dvgh-β gene expression did not change significantly ($P=0.65$) among the three feeding stages tested.

Pyrokinin-like receptor

A full length pyrokinin-like receptor was identified from a 2474 bp transcript with a 763 bp 5' UTR and a 139 bp 3' UTR. The gene encodes a 523-residue protein that contains a 7tm_1 motif. A BLASTp search with the protein sequence indicates the best match in the Genbank nr database is *Apis mellifera* pyrokinin-like receptor 2 (BLAST score, 391; E-value,

8e-107). A highly significant match by TBLASTn was also found in *I. scapularis* EST (EW782639; BLAST score, 419; E-value, 9e-117).

Expression in *D. variabilis* unfed female synganglia was 12.5-fold greater than in part-fed ticks (Tukey's test, $P<0.02$) (Figure 2). Expression levels were not significantly different between unfed and replete ticks ($P=0.09$) and part-fed and replete ticks ($P=0.49$).

Orcokinins

Four distinct orcokinin transcripts were identified in *D. variabilis*, of which only one is full length. The transcripts were putatively named orcokinin 1, 2, 3 and 4 (DvO1-4), in the order they were identified which does not imply physiological significance. Orcokinins 1 through 4 are 1052, 1136, 213 and 964 bp, respectively. DvO1 is 5' partial and contains a 553 bp 3' UTR, and encodes a 166-residue protein. DvO2 is the only full length orcokinin transcript identified. It contains a 65 bp 5' UTR and a 492 3' UTR. Translation of the transcript results in a 193-residue protein of which the first 18 amino acids are a predicted secretion signal. DvO3 lacks a start or stop codon and encodes a 70-residue fragment. DvO4 is a 5' partial transcript that contains a 700 bp 3' UTR. The coding region translates to an 88-residue peptide.

Christie (2008) previously showed the presence of ESTs that encode four orcokinins from *R. microplus*, two from *I. scapularis* and one from *A.americanum*. Comparison of these orcokinins to the same proteins in *D. variabilis* allowed for the prediction of the monobasic and dibasic cleavage sites (Figure 8). DvO1 contains a monobasic cleavage site where DvO2

contains a dibasic cleavage site due to the His^{28}/Arg^{58} conversion. Presence of Arg^{23} in DvO1 in the –6 position to the monobasic cleavage site conforms to the criteria proposed by Veenstra (2000). Comparison of DvO1, DvO2 and DvO4 indicates that DvO2 does not contain the NFDEIDRS peptide. DvO3 is the most divergent of the four sequences and contains three copies of NFDEIDRT(D/G)F(G/E)(G/E)F(R/Y), which is homologous to the typical arthropod orcokinin. The other three orcokinins contain two copies of this peptide and a third that contains an extended N-terminus.

Expression profiles of the four orcokinin genes were statistically the same (Figure 2). Levels in unfed ticks of DvO1, 2, 3 and 4 were 3.3-, 3.8-, 2.5- and 5.3-fold greater than part-fed ticks, respectively (Tukey's test, $P<0.004$, 0.0001, 0.0005 and 0.0001, respectively) and were 3.6-, 2.9-, 3.6- and 6.3-fold greater than replete ticks, respectively ($P<0.003$, 0.0001, 0.0001 and 0.0001, respectively). Expression of DvO1, 2, 3 and 4 were not significantly different ($P=0.98$, 0.19, 0.11 and 0.73, respectively).

Corazonin

A full length corazonin (ACC99624) was identified in *D. variabilis* that is a 545 bp transcript with a 90 bp 5' UTR and 191 bp 3' UTR. Translation results in an 88-residue peptide that contains a 21-residue secretion signal (Figure 9). Previously three separate corazonin encoding transcripts were identified (Christie, 2008) which are 53% identical in overlapping regions to *D. variabilis*. An alignment (Figure 9) demonstrates the conservation of the location of the secretion signal and dibasic and monobasic cleavage sites. As noted by

Christie (2008), the resulting peptide QTGQYSRGWTNG is predicted to have a cyclized N-terminal glutamine and amidated C-terminal glycine residue. All three of the *I. scapularis* corazonin encoding ESTs (EL516967, EL516402 and EL516698) are from a synganglion cDNA library. Only one corazonin homolog was located in the *D. variabilis* library in the current study.

qRT-PCR results showed that corazonin gene expression decreased with feeding to repletion only (Figure 2). Transcript levels in unfeds were not significantly different than part-fed female ticks (P=0.71); however there was a 2.4-fold decrease in replete ticks compared to unfed ticks (P<0.03). There was not a significant difference between part-fed and replete ticks (P<0.11).

Proprotein convertase (type2)

A full-length proprotein/prohormone processing protease was sequenced. The transcript is 4916 bp in length with a 2925 bp 3' UTR and a 29 bp 5' UTR. The resulting peptide sequence is 654 amino acids in length of which the first 21 residues are a predicted secretion signal. Two motifs are present in the protein. The first is a peptidase-S8 motif (PF00082) which spans residues 147 to 452. The sequence was also predicted to contain the catalytic triad, Asp^{173}, His^{214} and Ser^{389}, the order of which places this putative enzyme in the Subtilase family of serine proteases. The second motif in the protein is a proprotein convertase P-domain (PF01483) that is C-terminal to the peptidase-S8 motif and spans residues 511-611.

The sequence conservation of the putative *D. variabilis* type-2 proprotein convertase (DvPC2) with other proprotein convertases further supports the identification of this transcript. An alignment of DvPC2 with the same protein from other arthropod taxa demonstrates the high degree of relatedness of type-2 proprotein convertases (Supplementary Figure 1); overall these sequences are 50.8% identical. A neighbor-joining tree of the aligned proteins demonstrates that DvPC is most closely related to that of the crayfish, *Orconectes limosus* (AAK28328), as would be expected based on taxonomic position (Figure 10).

Since gene expression of DvPC in the tick synganglion may provide some information on the regulation of cleavage of prohormones to their active forms, qRT-PCR was used to monitor DvPC2 transcription. Expression levels in unfed female tick synganglion were 16.7-fold greater than part-fed ticks (Tukey's test, $P<0.006$) but were not significantly higher than replete ticks ($P=0.32$). It appeared that levels in replete ticks were higher than part-fed ticks however the difference was not significant ($P=0.06$). As with all of the qRT-PCR data in the current study, we did not determine whether the mRNA is translated at these time points in the synganglia.

Discussion

Reproduction and insulin signaling

Insect juvenile hormones (JH) and ecdysone are key regulators of larval development and reproduction in insects (Gilbert et al., 2000). Allatostatins (AST) were originally shown

to inhibit the synthesis of juvenile hormone by the corpora allata in insects *in vitro* (Rankin and Stay, 1987) while the allatotropins positively regulate JH biosynthesis (Kataoka et al., 1989). AST homologs have also been identified in crustaceans and are produced by the mandibular organs but their role in inhibition of methyl farnesoate synthesis in adults remains unclear (Kwok et al. 2005).

Three types of allatostatins have been characterized: FGLamides, W(X_6)Wamides and PISCFs (Stay and Tobe, 2007). Type-A allatostatins or FGLamides, so named for their C-terminal residues, have been shown to be allatostatic in cockroaches (Woodhead et al. 1989), crickets (Lorenz et al. 1995), and termites (Yagi et al. 2005). Data in a recent study in which the AST genes were silenced suggest that ASTs in *Spodoptera frugiperda* may be allatostatic (Meyering-Vos et al. 2006). The type-B and C ASTs have also been shown to be allatostatic in multiple insect species but with varying activity (Stay and Tobe, 2007). The AST sequenced in this study from *D. variabilis* and those located in *I. scapularis*, *R. microplus* and *L. laeta* all are type-A ASTs.

Previous to the current study, no allatostatin had been identified in the Chelicerata. Zhu and Oliver (2001) provided the first evidence of allatostatin activity in the American dog tick synganglion. Using an antibody against *Diploptera punctata* allatostatin, immunoreactive cells were located throughout the synganglion with the exception of little to no immunoreactivity in the pedal ganglia. The authors also reported that the strongest immunoreactivity was found in one month old unfed, virgin females suggesting that the majority of allatostatin-like peptides were produced prior to host attachment and blood

feeding. Our qRT-PCR results agree with these findings (Figure 2) and our results suggest that the protein is translated when the AST mRNA is present. Based on this finding, it is enticing to hypothesize that ASTs in ticks may regulate steps in the mevalonate pathway. In insects, the female gonadotropic hormone that initiates vitellogenesis is JH (Gilbert et al. 2000), with the exception of mosquitoes where this role is filled by ecdysteroids (Shapiro et al. 1986). In ticks, ecdysone is clearly the initiator of vitellogenesis and not JH (Neese et al., 2000; Thompson et al., 2005; Roe et al. 2008), although these studies are based on only one tick species in the former and one hard tick and one soft tick species in the latter. Additional studies of the role of ASTs in several tick species will be required to fully understand if a JH precursor is somehow involved in reproduction.

Insulin-like peptides (ILP) and insulin-like growth factors (ILGF) have been studied in a wide range of invertebrate taxa and have been shown to possess multiple functions such as control of lifespan (Corona et al., 2007; Kang et al, 2008), growth and metabolism (Rulifson et al., 2002) and reproduction (Nagasawa et al. 1984). An ILP from *Bombyx mori* was the first ILP demonstrated to have prothoraciotropic activity and was termed bombyxin or "small PTTH". The finding that insulin-like peptides are present in the tick synganglion is not surprising given the level of conservation of these peptides in the Arthropoda (Riehle et al. 2006) and the reports of insulin-like immunoreactivity in *D. variabilis* (Davis et al., 1994) and the soft tick, *O. parkeri* (Zhu and Oliver, 1991). Our results confirm that ticks do express an insulin-like peptide in the synganglion that is structurally conserved in terms of the location of the B-, C- and A- chains prior to cleavage and where the location of the intron

in the C-chain is similar to that of insects (Figure 3B) (Riehle et al. 2006). Zhu and Oliver (1991) found that unfed female tick synganglia contained more immunoreactive substance than did fed female ticks. qRT-PCR of the *D. variabilis* insulin-like peptide (DvILP) indicates that the message for this gene follows the same pattern, with higher levels in unfed female tick synganglion than blood-fed female ticks (Figure 2). The correlation of these data may suggest that tick ILP may not be involved in stimulating ecdysteroidogenesis unless further regulation of the hormone precedes release into the hemocoel. The alignment of tick ILP (Figure 3C) clearly indicates that tick ILPs are processed in a manner like that of insects, and this processing may serve as a regulatory step in ILP bioactivity. Other peptides involved in insulin signaling were identified from the current sequencing project, such as insulin-receptor substrate (FJ377313), indicating that at least part of this pathway is potentially active in adult ticks.

Feeding, diuresis and myotropic related peptides

Several hormones and hormone receptors that have been shown to be involved in feeding and diuresis were identified and include an ion-transport peptide, sulfakinin, sulfakinin receptor, leucokinin-like receptor and a calcitonin-like receptor. All of these except for the calcitonin-like receptor followed the same expression pattern (Figure 2) with significantly elevated levels in unfed female tick synganglia. Only the calcitonin-like receptor was significantly upregulated in part-fed and replete ticks.

ITPs have been shown to play a role is osmoregulation while closely related sequences such as CHHs can have different physiological functions. DvITP contains the CHH motif but the phylogenetic separation of CHHs and ITPs (Figure 5B) indicates that DvITP and IsITP are more closely related to ITPs than CHHs. The inclusion of MIHs to this analysis only further supports the hypothesis that the two tick sequences are more closely related to ITPs (data not shown). The branch point of DvIPT and IsITP occurs close to the base of the tree which may be an indication of multifunctionality of this peptide, which has been shown for other proteins in the CHH family (Chang et al., 1990; Liu et al. 1997). ITP in *Schistocerca gregaria* has been shown to control ion and fluid transport across the ileum *in vitro* (Audsley et al., 1992). Audsley et al. (2006) found that ITP-like peptide in *Schistocera gregaria* was upregulated during feeding. Tiu et al. (2007) reported on the osmoregulatory role in the Pacific white shrimp, *Litopenaeus vannamei* and the injection of 20-HE resulted in an increase in the expression of ITP. DvITP levels in the synganglion follow an opposite trend and decreased with blood-feeding and prior to mating (Figure 2). This finding suggests that ITP expression in ticks may not be regulated by the presence of 20-HE since significant levels would not be present until the tick has mated (Dees et al. 1984). These data also suggest that ITP does not function as a VIH since 20-HE levels do not rise in part-fed virgin ticks when the DvITP gene has been downregulated.

Sulfakinins are a group of neuropeptides related to cholecystokinin and gastrin that were first described in arthropods from the cockroach, *Leucophaea maderae* (Nachman et al. 1986). Sulfakinins increase the frequency of contractions in the cockroach hindgut, and

sulfation is required for this action in cockroaches. Nichols (2007) reported that this action can also occur in the *Drosophila melanogaster* larval anterior midgut and adult foregut with or without sulfated tyrosine residues and indicated this could be due to binding to different receptors. *D. variabilis* sulfakinin is predicted to be sulfated in the mature form and therefore may interact with specific GPCRs. Maestro et al. (2001) reported that this hormone inhibits food intake in cockroaches which has also been reported in the locust *Shistocerca gregaria* and is not surprising given the homology to cholecystokinins which functions as satiety inducing peptides (Lee et al. 1994.).

Sulfakinin and the putative sulfakinin receptor followed similar expression patterns in the synganglia of female ticks (Figure 2). Both decreased from unfed to part-fed ticks and remained low in repletes. Meyering-Vos and Müller (2007) found that sulfakinin expression was found only in the brain of the cricket, *Gryllus bimaculatus*, and decreased with age in adults. Since *D. variabilis* sulfakinin and the receptor both decreased with the initiation of blood-feeding, this may indicate a functional role of sulfakinins in feeding. These data, however, suggest that it is not acting as an indicator of satiety.

A leucokinin-like peptide receptor has been previously sequenced in the Southern cattle tick, *R. microplus* (Holmes et al. 2000). The same receptor presented here is only the second such sequence reported in the Chelicerata and one of only a few in the Arthropoda. In insects, leucokinins induce diuresis in the malpighian tubules and muscle contraction or myotropic effects on the hindgut (Gade, 2004; Howarth et al., 2002). Radford et al. (2004) cloned and sequenced *Anopheles gambiae* leucokinins and a leucokinin receptor. They

showed that like the tick leucokinin-like receptors, the mosquito leucokinin receptor is a GPCR. Additionally, they demonstrated that *in vitro*, leucokinins were the ligand to the cloned receptor and this binding raised intracellular calcium levels in a dose-dependent manner. They also showed that the receptor was expressed in stellate cells in the malpighian tubules. Our results indicate expression of the *D. variabilis* leucokinin-like receptor in the synganglion with the highest expression levels found in unfed female ticks. Tissue sources for this receptor in ticks other than the synganglion were not studied. Such studies could provide additional information on leucokinin signaling in the Acari.

Several genes typically associated with ecdysis behavior in insects were found in the adult library; they include a calcitonin-like receptor, eclosion hormone, bursicon and corazonin. The insect tanning pathway begins with decreasing 20-HE levels which leads to the release of ecdysis-triggering hormone (ETH) followed by eclosion hormone (EH) release and this positive-feedback results in pre-ecdysis behavior (Davis et al., 2007). Kim et al. (2004) showed that in *Manduca sexta*, injection of corazonin into pharate larvae elicited the release of ETH, indicating that this hormone in addition to 20-HE is involved in triggering pre-ecdysis behavior. EH causes the release of crustacean cardioactive peptide (CCAP) and a shift from pre-ecdysis to ecdysis behavior. Eclosion leads to bursicon expression, binding to its receptor and eventual increase of the second messenger cAMP and finally post-eclosion behavior.

Mating initiates the phase of rapid engorgement in female ticks and within a relatively short period of time they must stretch and expand their cuticle (Sonenshine, 1993).

In adults ticks, tanning is not involved during this cuticle expansion period (Andersen and Roepstorff, 2005), so the discovery of neuropeptides that have been linked to ecdysis behavior and tanning is intriguing. Furthermore, Dvbur-β was upregulated during feeding while Dvbur-α was not which is surprising given that bursicon is a cysteine-knot heterodimeric protein. Glycoprotein hormones are also cysteine-knot proteins consisting of alpha and beta subunits that in our current study also showed dissimilar expression profiles (Figure 2). These results may be explained by multiple forms of these neuropeptides. Multiple bursicon subunits were not identified, but multiple glycoprotein hormone subunits were putatively identified in at least one tick species supporting the possibility of this hypothesis for cysteine knot proteins in ticks.

Multiple orcokinin transcripts were identified in the current study and the same has been shown in another tick species (Christie, 2008). The alignment of the translations of these transcripts (Figure 8) and given that they all contain distinct open reading frames is evidence that they are not an artifact of the assembly from pyrosequencing. Orcokinins in the brain of other arthropod species have also been demonstrated (Yasuda-Kamatani, 2000; Hofer et al. 2005). Hofer and Homberg (2006) recently showed that orcokinins were predominantly present in the accessory medulla of the cockroach, *Leucophaea maderae*, and injection of labeled orcokinins into this region resulted in changes in circadian locomotor activity in a dose-dependent manner. The accessory medulla is a small neuropil in the optic lobe that is known to control circadian rhythms in locomotor activity. The authors concluded

that in this insect, orcokinin-related peptides possess multiple roles in the cockroach circadian system.

Hormone processing to bioactive mature peptides

Peptidic hormones are generally expressed as preprohormones that require removal of the secretion signal and cleavage at monobasic and/or dibasic endoproteolytic cleavage sites to produce the mature bioactive form. Scission at these sites in neuropeptides is carried out by type-2 proprotein convertases. Therefore, regulation of these proteolytic events provides a means to regulate the timing of the bioactivity of neurohormones since prohormones are rarely as active as the mature processed peptides. In *Drosophila melanogaster*, the PC2 gene, *amontillado* (*amon*), is required for completed embryo and larval development (Rayburn et al., 2003). They found that *amon* mutants successfully completed early molting cycles including apolysis but were incapable of shedding the first instar cuticle during ecdysis. Additionally, the correlation to phenotypes of flies with ablated brain insulin-producing cells was described and the authors noted that insulin and several other hormones involved in the hormonal cascade during ecdysis contain dibasic endoproteolytic cleavage sites that would be substrates for PC2. The same is true for some of the hormones described here which in hypothetically would make PC2 an attractive target for tick control since knockdown or inhibition of a single enzyme could result in the disruption of the entire hormonal cascade.

In summary, multiple putative neuropeptides and neuropeptide receptors were sequenced in tandem by high throughput pyrosequencing of the female tick synganglia transcritome of the American dog tick. The homologs of these D.v. proteins were also found in the recently sequenced *Ixodes scapularis* genome or in other tick EST databases. Previous immunocytochemical detection methods used to determine allatostatin and insulin-like peptide activity in the synganglion agreed with qRT-PCR data in the current study. The transcription of the hormones and receptors reported here do not necessarily indicate the presence of the peptide. Delayed translation of these hormones in the synganglion may serve as a mechanism which allows for precise timing of neuropeptide secretion and transport to target tissues.

Multiple hormones and receptors were characterized that potentially play a significant role during blood feeding and mating. However, more detailed studies will be needed to fully understand their importance in tick development and reproduction. As demonstrated here, the genome sequence of *Ixodes scapularis* coupled with high throughput pyrosequencing technologies will continue to be an enormous resource in furthering our understanding of tick biology.

References

Altschul, S.F., Madden, T.L., Schäffer, 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, S.O., Roepstorff, P. 2005. The extensible alloscutal cuticle of the tick, *Ixodes ricinus*. Insect Biochem. Mol. Biol. 35: 1181-1188.

Audsley, N., McIntosh, C., Phillips, J.E. 1992. Actions of ion transport peptide from locust corpus cardiacum on several hidgut transport processes. J. Exp. Biol. 173: 285-288.

Audsley, N., Meredith, J., Phillips, J.E. 2006. Haemolymph levels of *Schistocera gregaria* ion transport peptide and ion transport-like peptide. Physiol. Entomol. 31: 154-163.

Beldade, P., Rudd, S., Gruber, J.D., Long, A.D. 2006. A wing expressed sequence tag resource for Bicyclus anyana butterflies, an evo-devo model. BMC Genomics 7: 130-146.

Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol. Biol. 340: 783-795.

Chang, E.S., Prestwich, G.D., Bruce, M.J. 1990. Amino acid sequence of a peptide with both molt-inhibiting and hyperglycemic activites in the lobster, *Homarus americanus*. Biochem. Biophys. Res. Commun. 171: 818-826.

Christie, A.E. (2008) Neuropeptide discovery in Ixodoidea: An *in silico* investigation using publicly accessible expressed sequence tags. Gen. Comp. Endcrinol. 157: 174-185.

Corona, M., Velarde, R.A., Remolina, S., Moran-Lauter, A., Wang, Y., Hughes, K.A., Robinson, G.E. 2007. Vitellogenin, juvenile hormone, insulin signaling and queen honey bee longevity. Proc. Natl. Acad. Sci. USA. 104: 7128-7133.

Davis, H.H., Dotson, E.M., Oliver, J.H. 1994. Localization of insulin-like immunoreactivity in the synganglion of nymphal and adult *Dermacentor variabilis* (Acari: Ixodidae). Exp. Appl. Acarol. 18: 111-122.

Davis, M.M., O'Keefe, S.L., Primrose, D.A., Hodgetts, R.B. 2007. A neuropeptide hormone cascade controls the precise onset of post-eclosion cuticular tanning in *Drosophila melanogaster*. Development 134: 4395-4404.

Dees, W.H., Sonenshine, D.E., Breidling, E. 1984. Ecdysteroids in the American dog tick, *Dermacentor variabilis* (Acari: Ixodidae), during different periods of tick development

Gade, G. 2004. Regulation of intermediary metabolism and water balance of insects by neuropeptides. Annu. Rev. Entmol. 49: 93-113.

Gilbert, L.I., Granger, N.A., Roe, R.M. 2000. The juvenile hormones: historical facts and speculations on future research directions. Insect Biochem. Mol. Biol. 39:617-644.

Hofer, S., Dircksen, H., Tollbäck, P., Homber, U. 2005. Novel insect orcokinins: characterization and neuronal distribution in the brains of selected dicondylian insects. J. Comp. Neurol. 490: 57-71.

Howard, C.J., Prince, R.I., Dyker, H., Losel, P.M., Seinsche, A, Osborne, R.H. 2002. Structure-activity relationship of contractile effects induced by helicokinins in the isolated gut of the lepidoopteran caterpillar *Spodoptera frugiperda*. J. Insect Physiol. 48: 75-82.

Kang, M.-A., Mott, T.M., Tapley, E.C., Lewis, E.E., Luckhart, S. 2008. Insulin regulates aging and oxidative stress in *Anopheles stephensi*. J. Exp. Biol. 211: 741-748.

Kataoka, H., Toschi, A., Li, J.P., Carney, R.L., Schooley, D.A., Kramer, S.J. 1989. Identification of an allatotropin from adult Manduca sexta. Science 243: 1481-1483.

Kim, Y.-J., Spalovaská-Valachová, I., Cho, K.-H., Zitnanova, I., Park, Y., Adams, M.E., Žitňan, D. 2004. Corazonin receptor signaling in ecdysis initiation. Proc. Natl. Acad. Sci. USA 101: 6704-6709.

Kwok, R., Zhang, J.R., Tobe, S.S. 2005. Regulation of methyl farnesoate production by mandibular organs in the crayfish, *Procambarus clarkia*: A possible role for allatostatins. J. Insect Physiol. 51: 367-378.

Holmes, S.P., He, H., Chen, A.C., Ivie, G.W., Pietrantonio, P.V. 2000. Cloning and transcriptional expression of a leucokinin-like peptide receptor from the Southern cattle tick, *Boophilus microplus* (Acari: Ixodidae).

Lee, M.C., Schiffman, S.S., Pappas, T.N. 1994. Role of neuropeptides in the regulation of feeding behavior: a review of cholecstokinin, bombesin, neuropeptide Y and galanin. Neurosci. Biobehav. Rev. 18: 313-323.

Livak, K.J., Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\,\Delta C}$T method. Methods 25:402-408.

Liu, L., Laufer, H., Wang, Y.J., Hayes, T. 1997. A neurohormone regulating both methyl farnesoate synthesis and glucose metabolism in a crustacean. Biochem. Biophys. Res. Commun. 237: 694-701.

Lorenz, M.W., Kellner, R., Hoffmann, K.H. 1995. Identification of two allatostatins from the cricket, *Gryllus bimaculatus* de Geer (Ensifera: Gryllidae); additional members of a family of neuropeptides inhibiting juvenile hormone biosynthesis. Regul. Peptides 57: 227-236.

Maestro, J.L., Aguilar, R. Pascual, N., Valero, M.-L., Piulachs, M.-D., Andreu, D., Navarro, I., Bellés, X. 2001. Screening of antifeedant activity in brain extracts led to the identification of sulfakinin as a satiety promoter in the German cockroach. Eur. J. Biochem. 268: 5824-5830.

Meyering-Vos, M., Merz, S., Sertkol, M., Hoffman, K.H. 2006. Functional analysis of the allatostatin-A type gene in the cricket *Gyllus bimaculatus* and the armyworm *Spodoptera frugiperda*. Insect Biochem. Mol. Biol. 36: 492-504.

Meyering-Vos, M., Müller, A. 2007. Structure of the sulfakinin cDNA and gene expression from the Mediterranean field cricket *Gyllus bimaculatus*. Insect Mol. Biol. 16: 445-454.

Mitchell, R.D., Ross, E., Osgood, C., Sonenshine, D.E., Donohue, K.V., Khalil, S.M.S., Thompson, D.M., Roe, R.M. 2007. Molecular characterization, tissue-specific expression and RNAi knockdown of the first vitellogenin receptor from a tick. Insect Biochem. Mol. Biol. 37: 375-388.

Nachman, R.J., Holman, G.M., Haddon, W.F., Ling, N. 1986. Leucosulfakinin, a sulfated insect neuropeptides with homology to gastrin and cholecystokinin. Science 234: 71-73.

Nagasawa, H., Kataoka, H., Isogai, A., Tamura, S., Suzuki, A., Ishizaki, H., Mizoguchi A., Fujiwara, Y., Suzuki, A. 1984. Amino-terminal amino acid sequence of the silkworm prothoracicotropic hormone: Homology with insulin. Science. 4680: 1344-1345.

Neese, P.A., Sonenshine, D.E., Kallapur, V.L., Apperson, C.S., Roe, R.M. 2000. Absence of juvenile hormones in the American dog tick, *Dermacentor variabilis* (Say) (Acari: Ixodidae) and *Ornithodoros parkeri* (Acari: Argasidae). J. Insect Physiol. 46: 477-490.

Radford, J.C., Terhzaz, S., Cabrero, P., Davies, S.-A., Dow, J.A.T. 2004. Functional characterisation of the *Anopheles* lecokinins and their cognate G-protein coupled receptor. J. Exp. Biol. 207: 4573-4586.

Rankin, S.M., Stay, B. 1987. Distribution of allatostatin in the adult cockroach, *Diploptera punctata* and effects on corpora allata *in vitro*. J. Insect Physiol. 33: 551-558.

Rayburn, L.Y.M., Gooding, H.C., Choksi, S.P., Maloney, D., Kidd, A.R., Siekhaus, D.E., Bender, M. 2003. *amontillado*, the drosophila homolog of the prohormone processing protease PC2, is required during embryogenesis and early laval development. Genetics 163: 227-237.

Riehle, M.A., Fan, Y., Cao, C., Brown, M.R. 2006. Molecular characterization of insulin-like peptides in the yellow fever mosquito, *Aedes aegypti*: Expression, cellular localization, and phylogeny. Peptides 27: 2547-2560.

Roe, R.M., Donohue, K.V., Khalil, S.M.S., Sonenshine, D.E. 2008. Hormonal regulation of metamorphosis and reproduction in ticks. Frontiers Biosci. 13: 7250-7268.

Rulifson, E.J., Kim, S.K., Nusse, R. 2002. Ablation of insulin-producing neurons in flies: growth and diabetic phenotypes. Science 296: 1118-1120.

SAS Institute. 2001. SAS/STAT, version 8.02. SAS Institute Inc., Cary NC, USA.

Shapiro, A.B., Wheelock, G.D., Hagedorn, H.H., Baker, F.D., Tsai, L.W., Schooley, D.A. 1986. Juvenile hormone and juvenile hormone esterase in adult females of the mosquito *Aedes aegypti*. J. Insect Physiol. 32: 867-877.

Seixas, A., Freisen, K.J., Kaufman, K.R. 2008. Effect of 20-hydroxyecdysone and haemolymph on oogenesis in the ixodid tick *Amblyomma hebraeum*. J. Insect Physiol. 54: 1175-1183.

Sonenshine, D.E. 1993. Biology of Ticks vol. 2. Oxford University Press, NY.

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

Thompson, D.M., Khalil, S.M.S., Jeffers, L.A., Ananthapadmanaban, U., Sonenshine, D.E., Mitchell, R.D., Osgood, C.J., Apperson, C.S., Roe, R.M. 2005. In vivo role of 20-hydroxyecdysone in the regulation of the vitellogenin mRNA and egg development in the American dog tick, *Dermacentor variabilis* (Say). J. Insect Physiol. 51: 1105-1116.

Woodhead, A.P., Stay, B., Seidel S.L., Khan, M.A., Tobe, S.S. 1989. Primary structure of four allatostatins: neuropeptide inhibitors of juvenile hormone synthesis. Proc. Natl. Acad. Sci. USA 86: 5997-6001.

Wei, Z., Baggerman, G., Nachman, R.J., Goldsworthy, G., Verhaert, P., De Loof, A., Schoofs, L. 2000. Sulfakinins reduce food intake in the desert locust, *Schistocerca gregaria*. J. Insect Physiol. 46: 1259-1265.

Yagi, K.J., Elliott, K.L., Teesch, L., Tobe, S.S., Stay, B. 2008. Isolation of cockroach Phe-Gly-Leu-amide allatostatins from the termite *Reticulitermes flavipes* and their effect on juvenile hormone synthesis. J. Insect Physiol. 54: 939-948.

Yasuda-Kamatani, Y., Yasuda, A. 2000. Identification of orcokinin gene-related peptides in the brain of the crayfish *Procambarus clarkia* by the combination of MALDI-TOF and on-line capillary HPLC/Q-Tof mass spectrometries and molecular cloning. Gen. Comp. Endocrinol. 118: 161-172.

Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R., Siebert, P.D. 2001. Reverse transcriptase template switching: A SMART approach for full-length cDNA library construction. Biotechniques 30: 892-897.

Zhu, X.X., Oliver, J.H. 1991. Immunocytochemical localization of an insulin-like substance in the synganglion of the tick *Ornithodoros parkeri*. 13: 153-159.

Zhu, X.X., Oliver, J.H. 2001. Cockroach allatostatin-like immunoreactivity in the synganglion of the American dog tick *Dermacentor variabilis* (Acari: Ixodidae).

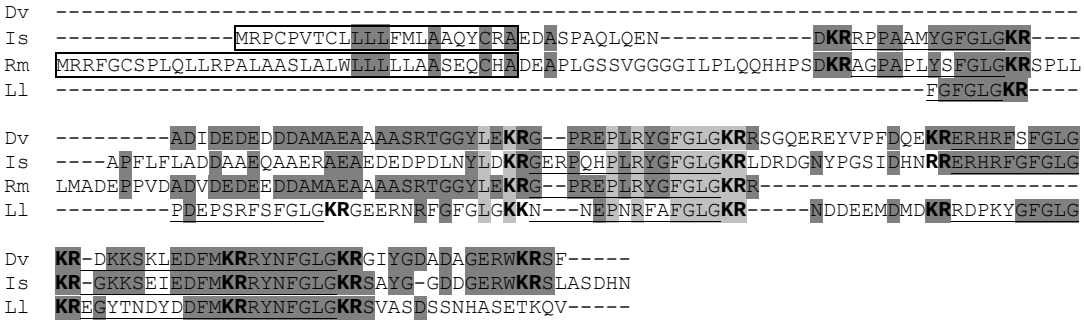


Figure 1. Multiple isoforms of allatostatin type-A are conserved in the Chelicerata. Boxed residues are predicted secretion signals, residues in bold indicate predicted putative dibasic endoproteolytic cleavage sites based on comparison with other known type-A allatostatins, residues with underline indicate predicted isoforms of mature allatostatins prior to amidation at the carboxy terminus. *Dermacentor variabilis* (Dv; ACC99603), *Ixodes scapularis* (Is; ABJB010344865), *Rhipicephalus microplus* (Rm; CK192037), *Loxoceles laeta* (Ll; EY189417). Light grey shading represents identical residues and dark grey is conserved residues.

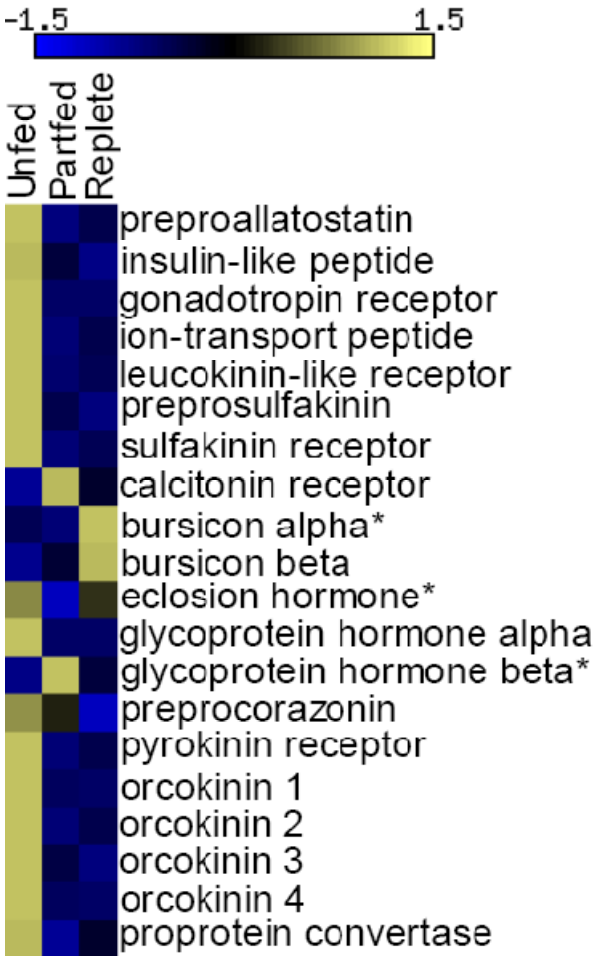
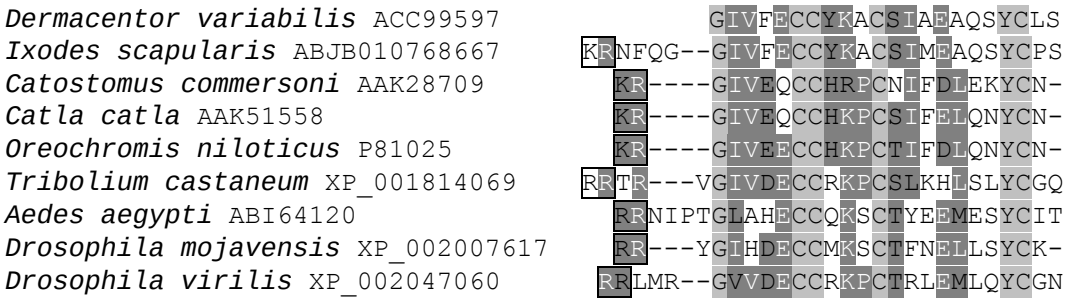
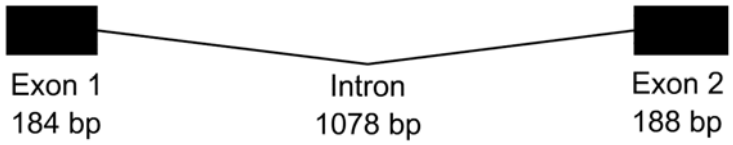


Figure 2. Gene expression of putative neurohormones, receptors and proprotein convertase in the synganglion of the American dog tick, *Dermacentor variabilis* during blood-feeding and mating. Fold expression levels within each gene were normalized to a mean of zero and a variance of one to aid in visualization of the data. Genes marked with an asterisk did not have significant changes in expression. Scale bar represents normalized fold changes in expression, yellow represents increased expression, black represents no change in expression and blue indicates decreased expression within each transcript. Non-transformed fold gene expression levels and associated ANOVA test statistics are available in Supplementary Table 1.

A



B



C

```
Secretion Signal: (1) MVSWALNTVVVALVAASALVAPAAA
B-chain:          (26)GSGRRCGKILLEFMEFVCEGEFYDPYENTGP
C-chain:          (57)KRSLIGQRLFPLVSPGIENTDKAPASGFLRAESASQLLR
A-chain:          (96)KRNFQGGIVFECCYKACSIMEAQSYCPS
```

Figure 3. Insulin-like peptides (ILP) in ticks are structurally conserved within the Arthropoda. A. Alignment of sequence data from *Dermacentor variabilis* with *Ixodes scapularis* (predicted) and other ILP alpha-chains identified by BLASTp. Boxed residues indicate predicted dibasic endoproteolytic cleavage sites. Light grey residues represent identical residues and dark grey are conserved residues. B. Intron-exon structure of *I. scapularis* ILP. C. Complete predicted protein sequence of *I. scapularis* ILP separated by secretion signal, B-chain, C-chain and A-chain. Boxed guanine indicates the location of the intron shown in B.

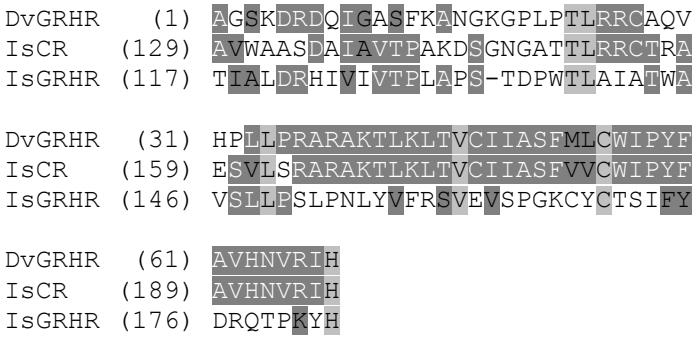


Figure 4. Alignment of *Dermacentor variabilis* putative gonadotropin releasing hormone receptor (DvGRHR) with translations of *Ixodes scapularis* predicted cardioacceleratory peptide receptor (IsCR)(DS713552) and *I. scapularis* EST EW846470 putatively identified as a GRHR (IsGRHR). DvGRHR is 58.8% identical to IsCR and 16.2% identical to IsGRHR in the aligned region. Light grey shading represents identical residues, dark grey/white text are conserved residues, dark grey/black text are similar residues.

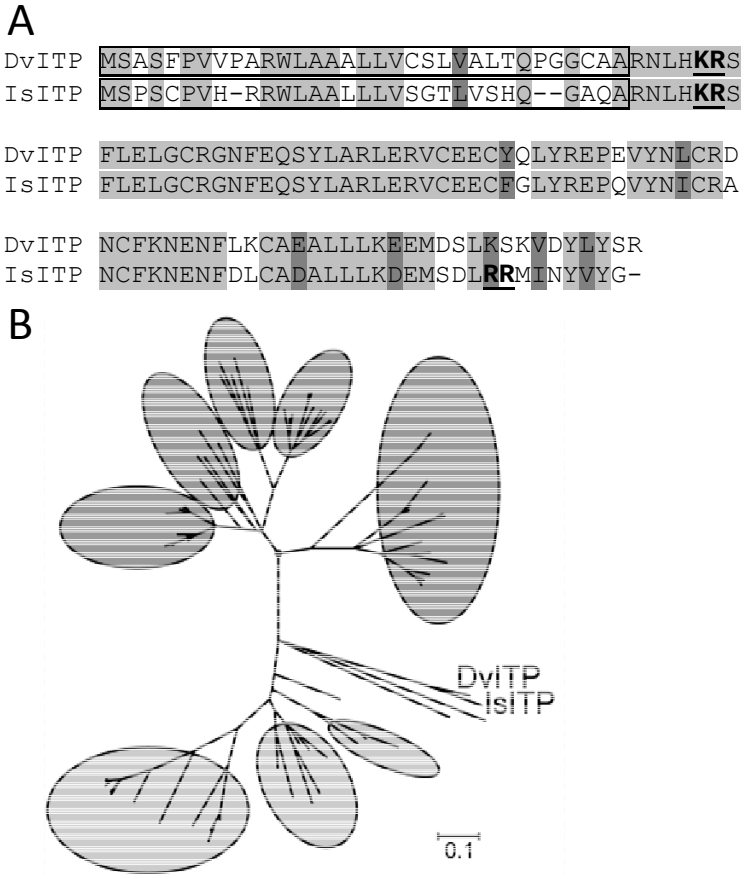


Figure 5. Ion-transport peptides in ticks are conserved and more similar to other ion-transport peptides than crustacean hyperglycemic hormones. A. Alignment of *Dermacentor variabilis* ion-transport-like peptide (ACC99599; 114 aa) with putative *Ixodes scapularis* ion-transport-like peptide (EW937910; 110 aa). Boxed residues represent the predicted secretion signal and bold and underlined residues indicate predicted endoproteolytic cleavage sites. Light grey shading indicates identical residues and dark grey indicates conserved residues. The sequences are 68.4% identical. B. Neighbor-joining tree of DvITP and IsITP with ITPs (light grey circles) and CHHs (dark grey circles). Included sequences were retrieved from the nr Genbank database from the top 100 hits by BLASTp. The optimal tree with the sum of branch length = 10.4 is shown. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site.

A

```
Dv  (1) MQLPARFLLFLLVAIMAAASSALGYSAASP
Is  (1) MRASSWFLLCLLAALVYGSWS---------

Dv (31) VSSQQQQQQRHRISVGRWLKSVLPAAAAAA
Is (20) --SPASLQQHHRMAMGKWLKSVLPGAP---

Dv (61) SAGDSDSRNTADLDAADMIDPVLLASGFAK
Is (47) SGGDAGSRNSGDIDT-DMIDPVILANGFAK

Dv (91) RQEDDYGHMRFGRSDDYGHMRFGRK
Is (76) RQDDDYGHMRFGRSDDYGHMRFGRK
```

B

```
DvSR  (1) -GRAPKNAKGRQFIFRGNYRKSRASKRRVI
IsSR  (1) VELRRQKLNRRPVILRGNYKKSRASKRRVI

DvSR (30) RMLSVLVLEFFVCWTPLYVLHTWTVFDAHA
IsSR (31) RMLSVLVLEFFVCWMPLYVLHTWAVFDAES

DvSR (60) AYSRVPAGAFAAVHLL--------------
IsSR (61) AYSHVSASGIAAVHLLAYVSSCCNPITYCF
```

Figure 6. Sulfakinin and sulfakinin receptors present in ticks. A. Alignment of *Dermacentor variabilis* sulfakinin (Dv) (ACC99604) with *Ixodes scapularis* putative sulfakinin (Is) (EW941557). The sequences are 57.4% identical. Boxed residues are predicted secretion signal; residues in bold and underline are predicted dibasic and monobasic cleavage sites. B. Alignment of *D. variabilis* sulfakinin receptor (DvSR) and translation of *I. scapularis* EST EW798836 (IsSR). The sequences are 70.7% identical in the aligned regions above. Light grey shading indicates identical residues and dark grey are similar residues.

A

```
Is bur-α   (1) MLICVRSPCSLLASWLILAVLAASMGPEES

Dv bur-α   (1) ------------------------------
Is bur-α  (31) CQLRPVIHVLKQPGCQPKPIPSFACQGSCS
                             * **
Dv bur-α   (1) -----SGSRYWQVERSCMCCQEMGEREATK
Is bur-α  (61) SYVQVSGSRYWQVERSCMCCQEMGEREATK
                  *                  *    *
Dv bur-α  (26) AVFCPKGPGPKFRKLVTRAPVECMCRPCTA
Is bur-α  (91) AVFCPKGPGPKFRKLITRAPVECMCRPCTA

Dv bur-α  (56) PDEASVLPQEFVGL
Is bur-α (121) PDEASILPQEFVGL
```

B

```
Dv bur-β   (1) MTWRALTVAAAISALWLTTAAVSASLASAL
Is bur-β   (1) ------MKCWTCGVLWLVLLARPRATLDGG
                        *
Dv bur-β  (31) EGPGGGVASCRLQETSIRITRDHSDDQGSP
Is bur-β  (25) -------SSCRLQPTTIRITRDQNDELGSL
                  *         *   *
Dv bur-β  (61) VRTCEGTVLVSRCEGTCVSQVQPSITLPHG
Is bur-β  (48) TRTCEGTVLVSRCEGTCISQVQPSITLPHG

Dv bur-β  (91) FLK--------------------------
Is bur-β  (78) FLKECNCCRETYMNKREIQLQDCFDPNGQK

Dv bur-β  (94) -------------------------
Is bur-β (108) LYGAEGSMTIFLEEPQDCSCHKCGG
```

Figure 7. Bursicon subunits, bur-α and bur-β are encoded by separate transcripts. A. Alignment of *Dermacentor variabilis* (Dv) and *Ixodes scapularis* (Is) bur-α. The sequences are 97.1% identical in the region of overlap. B. Alignment of Dv and Is bur-β. The sequences are 67.8% identical in the overlapping regions of the mature peptides. Boxed residues are predicted secretion signals; conserved cysteine residues are indicated with asterisks. Light grey shading represents identical residues and dark grey are similar residues.

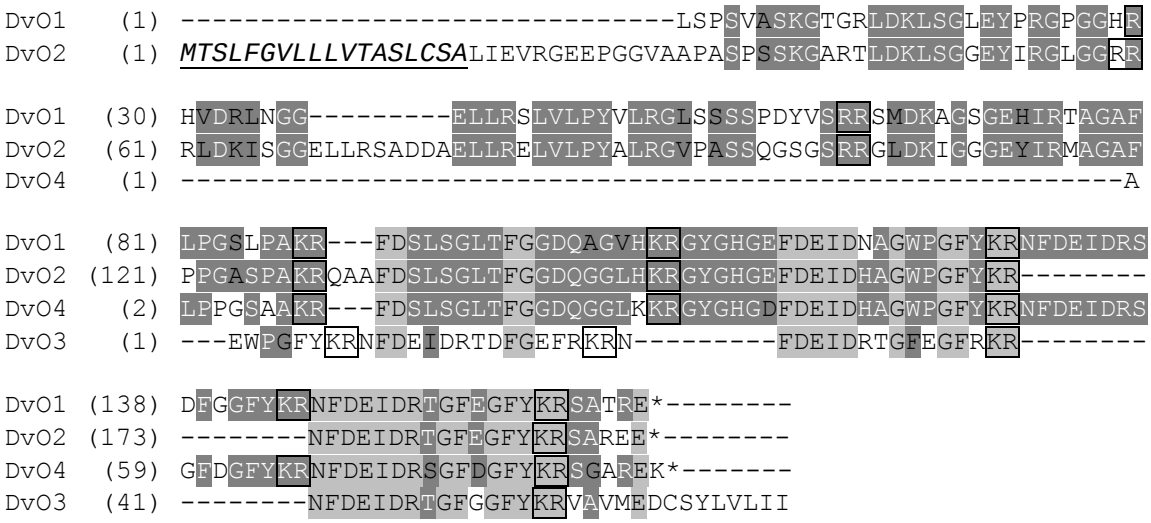


Figure 8. Alignment and proposed cleavage of *Dermacentor variabilis* orcokinins (DvO). The secretion signal in DvO2 is shown in italics and underlined. Predicted monobasic and dibasic cleavages sites are boxed. Light grey background represent identical residues, dark grey background/black text are similar residues and dark grey background/white text are conserved. Asterisks indicate the location of the stop codon.

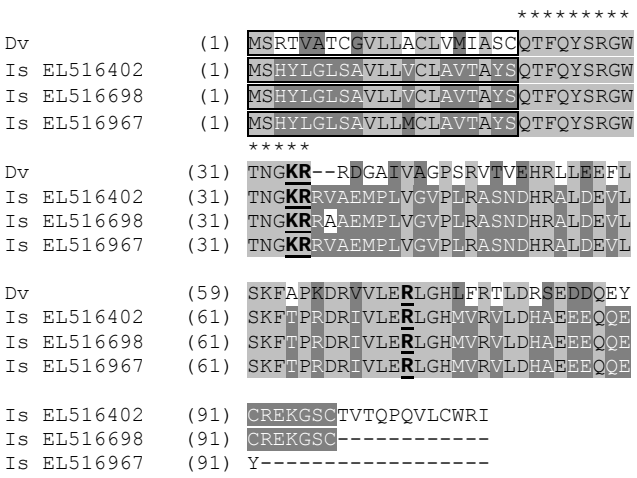


Figure 9. Alignment of *Dermacentor variabilis* (Dv) putative corazonin and *Ixodes scapularis* (Is) translations of ESTs EL516402, EL516698, EL516967. Boxed residues indicate predicted secretion signal and bold/underline residues are predicted dibasic and monobasic endoproteolytic cleavage sites. The resulting mature peptide (below asterisks) are absolutely conserved in corazonins studied thus far in insects. Light grey shading represents identical residues, dark grey background/black text are similar residues and dark grey background/white text are conserved.

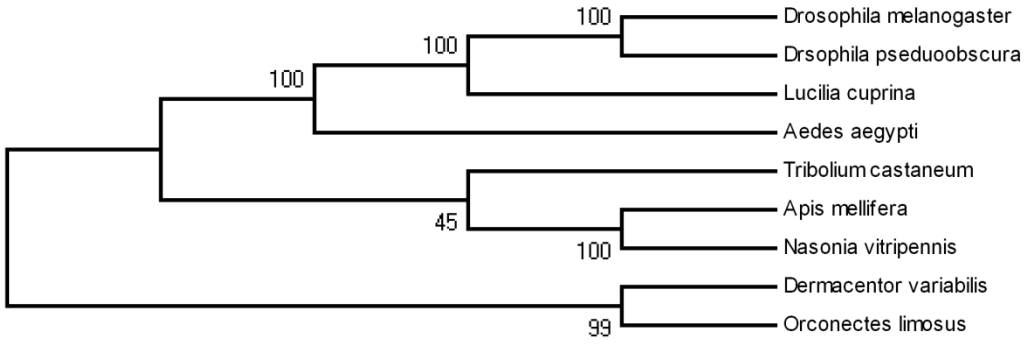
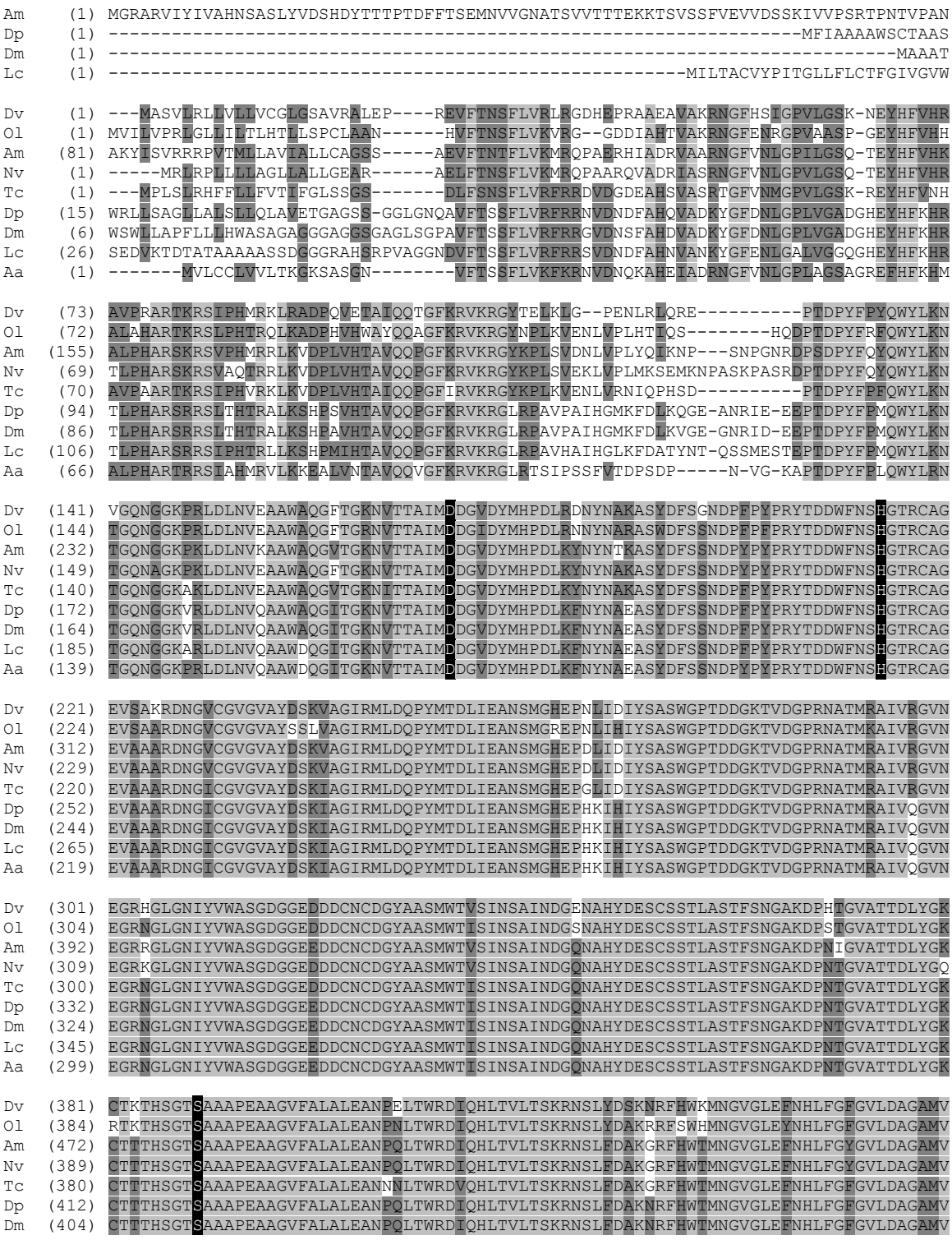
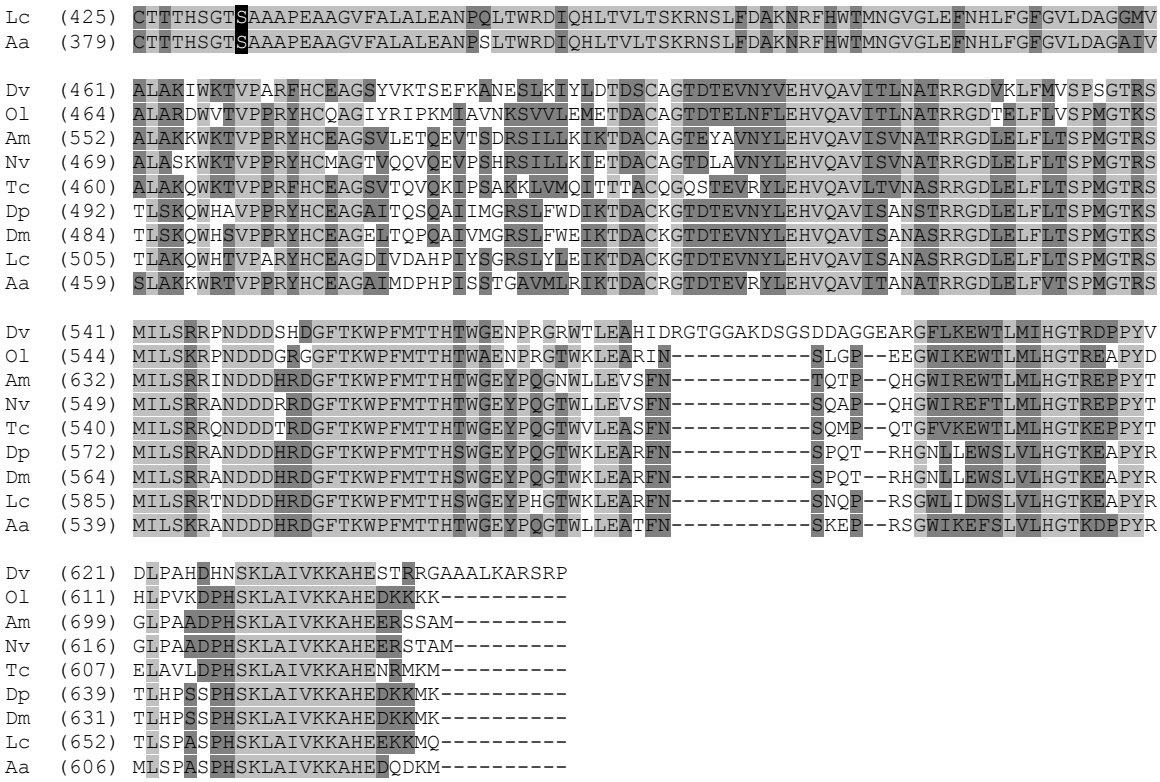


Figure 10. Type 2 proprotein convertase in *Dermacentor variabilis* is related to type 2 proprotein converstases in other arthropod taxa. The optimal neighbor-joining tree with the sum of branch length = 0.85 is shown. Accession numbers are: *Dermacentor variabilis* (ACD63025), Orconectes limosus (AAK28328), Apis mellifera (XP_392366), Nasonia vitripennis (XP_001600872), Tribolium castaneum (XP_972593), Drosophila pseudoobscura (XP_001358444), Lucilia curpina (CAA70106), and Aedes aegypti (XP_001648949). Taxa were identified by BLASTp and had an expect value of 0.00 compared to Dv. An alignment of the sequences is shown in Supplementary Figure 1.

Appendix A. Supplementary material





Supplementary Figure 1. Alignment of *Dermacentor variabilis* (Dv) proprotein convertase (type 2) with the translated sequence from Orconectes limosus (Ol; crayfish; AAK28328), Apis mellifera (Am; XP_392366), Nasonia vitripennis (Nv; parasitoid wasp; XP_001600872), Tribolium castaneum (Tc; XP_972593), Drosophila pseudoobscura (Dm; XP_001358444), Lucilia cuprina (Lc; CAA70106), and Aedes aegypti (Aa; XP_001648949). Taxa were identified by BLASTp and had an expect value of 0.00 compared to Dv. Light grey shading represents identical residues and dark grey represents conserved residues. The conserved Asp/His/Ser catalytic triad is shown with white text and black background.

Supplementary Table 1. Fold gene expression of putative neurohormones, receptors and proprotein convertase in the synganglion of the American dog tick, *Dermacentor variabilis* during blood-feeding and mating. Values followed by the same letter within each gene are not statistically significantly different. Pairwise-comparisons were performed only for those genes that were showed a significant difference in the analysis of variance. Within each gene, values followed by the same letter are not statistically different.

	Fold Gene Expression				
Gene	Unfed	Part-fed	Replete	F-value	p-value
Bursicon Alpha	1.00	0.96	1.33	0.58	0.58
Bursicon Beta	1.00a	4.22b	12.80c	62.4	0.0001
Calcitonin receptor-like Receptor	1.00a	2.28b	1.41b	4.82	0.04
Eclosion Hormone	1.00	0.09	0.76	2.49	0.14
Glycoprotein Hormone Alpha	1.00a	0.23b	0.22b	119.4	0.0001
Glycoprotein Hormone Beta	1.00	1.66	1.15	0.45	0.65
Gonadotropin Receptor	1.00a	0.31b	0.32b	6.81	0.016
Insulin-like peptide	1.00a	0.37b	0.17b	15.73	0.001
Ion-transport peptide	1.00a	0.26b	0.35b	8.59	0.008
Leucokinin-like receptor	1.00a	0.06b	0.13b	12.22	0.003
Orcokinin 1	1.00a	0.3b	0.28b	14.36	0.002
Orcokinin 2	1.00a	0.26b	0.35b	50.13	0.0001
Orcokinin 3	1.00a	0.4b	0.28b	37.84	0.0001
Orcokinin 4	1.00a	0.19b	0.16b	45.44	0.0001
Preproallatostatin	1.00a	0.31b	0.41b	23.87	0.0003
Preprocorazonin	1.00a	0.8a,b	0.42b	5.11	0.03
Preprosulfakinin	1.00a	0.31b	0.18b	25.06	0.0002
Proprotein convertase	1.00a	0.06b	0.36a,b	8.80	0.007
Pyrokinin-like receptor	1.00a	0.08b	0.19a,b	6.53	0.02
Sulfakinin Receptor	1.00a	0.23b	0.31b	14.17	0.002