CHARACTERIZATION OF HEME LIPOPROTEIN IN IXODID TICK SALIVA AND HEMOLYMPH

By

WILLIAM JUSTIN CORDILL

Bachelor of Science in Biochemistry and Molecular Biology

Oklahoma State University

Stillwater, Oklahoma

2005

Submitted to the Faculty of the
Graduate College of
Oklahoma State University
in partial fulfillment of the
requirements for the Degree of
MASTER OF SCIENCE
July, 2007

CHARACTERIZATION OF HEME LIPOPROTEIN IN

IXODID TICK SALIVA AND HEMOLYMPH

Thesis Approved:

Jack W. Dillwith
Thesis Adviser

Haobo Jiang

Jose Soulages

A. Gordon Emslie
Dean of the Graduate College

ACKNOWLEDGMENTS

Leading up to and during this project, I have been blessed with an amazing support system of family, friends, and mentors. I want to express my deep gratitude to Dr. Jack Dillwith, my advisor and mentor, for all of his guidance, support, and time that he has invested into me. He is a great scientist and advisor, not to mention a great person. Many, many thanks to my other committee members, Dr. Haobo Jiang and Dr. Jose Soulages, for their patient instruction and time as I pursued my graduate studies. To Robin Madden, what can I say? She has taught me so much, has spent so much of her time doing it; has suffered my breaks and spills, my forgetfulness, and my messiness; and has endured being the butt of my jokes for over two years. Thank you.

Without the ever-present love, support, and encouragement of my parents, I would not be where I am today. Thank you! It is rare these days to find such wonderful parents as you. Similarly, I want to thank my sister and grandparents (Nana and Papa, Grandma and Grandpa) for their love and encouragement. Of all of my family, and of all people, most importantly: my beautiful wife, my best friend. Thank you for your love and encouragement and especially your patience.

Lastly, without the grace of God I would not love biology or science in general. Without the grace of God I could not do what I have done or what I will one day do. And without the grace of God I would not be alive. May this work glorify Him.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Ticks (Acari) are hematophagous ectoparasites whose economic and medical impacts on human welfare cost the world billions of dollars annually (Sonenshine 1993). Only mosquitoes rival ticks in the number of pathogens for which they act as vectors (Obenchain and Galun 1982). Hard ticks (Ixodidae) compose 80% of all tick species and especially contribute to the pest status of ticks. The importance of controlling Ixodid ticks justifies the investigation of Ixodid tick survival, or more specifically, Ixodid tick feeding mechanisms. Crucial to successful Ixodid tick feeding, which in some cases takes weeks on a host, is the ability of the slow feeding parasite to evade a host's detection and defenses (Bowman *et al* 1997). Tick saliva has been found to contain a number of factors that counteract the host's detection and defense mechanisms at a number of points or allow the tick to sequester essential nutrients from the host (Bowman *et al* 1997, Maya-Monteiro *et al* 2000).

A recently characterized protein in hard tick hemolymph named heme-lipoprotein (HeLp), also called carrier protein (Cp), is the most abundant protein in tick hemolymph (Maya-Monteiro *et al* 2000, Gudderra *et al* 2001). Recently it was found that this protein is also the most abundant protein in tick saliva and that it is highly conserved across Ixodid ticks (Madden *et al* 2002, Dillwith unpublished). The abundance of this protein suggests it serves a vital role in Ixodid tick feeding and thus survival. It was proposed that the properties of the HeLp protein in the lonestar tick, *Amblyomma americanum*, be further investigated for the elucidation of its

function. The lonestar tick, *Amblyomma americanum*, has been chosen for this study because of its availability and representation of the large and economically and medically significant Ixodid genus *Amblyomma* (Sonenshine 1991). Also, *Amblyomma americanum* itself is becoming more medically significant with the discovery of its status as vector of the bacterium causing human ehrlichioses (Childs and Paddock 2002).

CHAPTER 2

RESEARCH GOAL AND OBJECTIVES

The goal of this research was to provide information that will help reveal the structural and functional properties of Ixodid heme lipoprotein. This research had the three following specific objectives:

1. *To compare HeLp homologues among Ixodid tick species*. A combination of MALDI-TOF mass spectrometry analysis of HeLp tryptic digests and Edman degradation were employed. We hypothesized that HeLp, because of its abundance in the Ixodid ticks examined and its proposed role as a carrier of essential but toxic heme iron, is conserved among Ixodid ticks.
2. *To perform structural studies of hemolymph and salivary HeLp*. Reducing and non-reducing SDS-PAGE, immunoblotting, MALDI-TOF MS analysis, various staining techniques, density gradient ultracentrifugation, spectrophotometry, and lipid analysis were employed to reveal HeLp structure and discover HeLp-related proteins. We hypothesize, based on the interpretation of our preliminary data, that HeLp-B occurs as a dimer that associates with HeLp-A and that unprocessed HeLp is present in the hemolymph and saliva.
3. *To obtain HeLp transcript sequence information*. Various PCR and/or cloning techniques were employed for *A. americanum* HeLp and the derived sequences compared with other available HeLp sequences. We hypothesize *A. americanum* HeLp is transcribed as a

single 210 kDa peptide in manner consistent with the HeLp sequences from *Dermacentor variabilis* and *Rhipicephalus sanguineus*.

CHAPTER 3

REVIEW OF LITERATURE

Ticks as Pests and Disease Vectors

Ticks are hematophagous ectoparasites belonging to the Superfamily Ixodoidea of the Order Acari (Phylum Arthropoda, Class Arachnida). Three families compose the ticks: the Ixodidae (hard ticks), the Argasidae (soft ticks), and a third minor family, the monotypic Nutalliellidae (Sonenshine 1991). Only mosquitoes are vectors for more human diseases than ticks, and ticks are vectors of more kinds of pathogens than any other arthropod group (Obenchain and Galun 1982, Sonenshine 1991). Notably, ticks are vectors for the spirochete *Borrelia burgdorferi* which causes Lyme disease, the most common vector-borne human disease in the U.S. (Gale and Ringdahl 2001). Although not fatal, symptoms of Lyme disease are diverse and often serious, involving neurological and caridiac effects, among others (Walker 1998). Other tick-borne diseases include the human monocytic and granulocytic ehrlichioses, Rocky Mountain spotted fever, and the babesioses, all of which may be fatal (Walker 1998, Gale and Ringdahl 2001). Ticks cause billions of dollars in damage to livestock production worldwide, especially with cattle (Sonenshine 1993). The cost of healthcare and veterinary treatment due to the diseases spread by ticks to both humans and pets is unknown, and money spent on the control of ticks is an additional cost (Sonenshine 1993).

About 80% of ticks are hard (Ixodid) ticks (Jongejan and Uilenberg 2004). Hard ticks are thus named because their cuticles have larger sclerotized (hardened) regions than soft ticks

(Hackman and Filshie 1982). This hard cuticle is needed to protect the tick from host grooming because, in contrast to the fast-feeding Argasids, Ixodid ticks require prolonged feeding periods (up to weeks) (Mans and Neitz 2004).

The Ixodid Tick Lifecycle

Sonenshine (1991) gives a description of the Ixodid tick lifecycle: Ixodid ticks have four lifestages—the embryonated egg, the larva, the nymph, and the adult. The lonestar tick, like most Ixodid ticks, has a 3-host life cycle—that is, at each of the latter 3 lifestages, the tick finds a host, attaches, feeds to repletion (unless it is an adult male), falls off the host, and, if not an adult, molts to the next lifestage. The tick then attaches to another host, although not necessarily the same species as the previous host. Attachment and feeding is a slow process, taking days for Ixodid larvae and nymphs and even up to two weeks for adult females to become engorged. Adult males do not feed to repletion; rather, they feed intermittently on one host and mate with the female while she is feeding. Adult females feed to repletion, often reaching sizes 100 times that of unfed adult females. Most of her body weight is converted to the thousands of eggs she will lay after she has dropped off the host. After oviposition, the female dies.

Tick Saliva

Although the adult female ultimately gains an enormous mass and volume through her bloodmeal, much of the blood she imbibes is excreted back into the host through the salivary glands (Saur *et al* 2000). Because of the prolonged periods of attachment, and because of the exchange of both tick and host fluids between host and parasite, Ixodid tick saliva contains a number of factors that counteract the host's immune, inflammatory, and haemostatic defenses

(Bowman *et al* 1997). Central to the production and excretion of these factors are the salivary glands (Saur *et al* 1995).

The salivary glands consist of a pair of branched, clustered alveoli (or acini) that originate anterolaterally and extend posteriorly within the lateral sides of the tick (Saur *et al* 1995). Female Ixodidae have one type of ungranulated acinus (Type I acini) and two types of granulated acini (Types II and III acini) (Sonenshine 1991). Each type of acinus contains morphologically distinct cell types (Sonenshine 1991). Type III acini are the most abundant in the salivary glands, and they contain the f-cells, which undergo a substantial physiological and morphological change during feeding (Sonenshine 1991). However, all of the acini enlarge during feeding, though the number of cells does not change (Saur *et al* 1995). The acini arise directly or indirectly from both of the central salivary ducts which originate anteriorly from the single salivarum and extend posteriorly (Sonenshine 1991). The salivarum is connected to the buccal cavity, through which both host and tick fluid pass at different times in alternating directions (Saur *et al* 1995).

An excellent review (Bowman *et al* 1997) describes the known factors in tick saliva and tick salivary gland extract that thwart the host's haemostatic, inflammatory, and immune responses: Ticks deliver anti-haemostatic factors into the host tissue that inhibit host platelet adhesion, activation and aggregation; inhibit blood coagulation; and inhibit vasoconstriction while promoting vasodilation (Bowman *et al.* 1997). Apyrase, an enzyme that catalyzes the hydrolysis of ADP which is released from damaged cells and activates platelets, has been found in Ixodid ticks, though it is apparently absent in *A. americanum* (Ribero *et al.* 1987, Bowman *et al.* 1997). A number of small peptides (less than 20 kDa) in soft ticks, including Tick Adhesion Inhibitor, prevents the aggregation or adhesion of platelets (Bowman *et al.* 1987, Karczewski *et*

al. 1995). Prostaglandins have been found in salivary secretions that prevent platelet aggregation (Bowman *et al.* 1996, 1997). Prostacyclin especially inhibits aggregations and even disaggregates aggregated platelets (Radomski *et al.* 1987). Prostaglandins from at least some Ixodid tick species also play a role in vasodilatation (Bowman *et al.* 1997). A number of anticoagulant factors exist in tick saliva that directly or indirectly prevents the thrombin induced conversion of soluble fibrinogen into insoluble fibrin (Bowman *et al.* 1997). Most of these peptides are less than 20 kDa. Though these factors are important in preventing coagulation, they play the important role of inhibiting thrombin's more potent ability to cause platelet aggregation (Packham 1994).

Bowman *et al*. (1997) continue the review: For successful feeding ticks must also overcome or suppress both the immune and inflammatory responses. Although little is known about the factors responsible for this, saliva and salivary gland extracts show immunosuppressive effects *in vitro*, including the suppression of complement, antibody production, cytokine production, and T-lymphocyte response (Wikel *et al.* 1994, Bowman *et al.* 1997). Both peptides and lipid soluble factors have been found to inhibit these responses (Bowman *et al.* 1997). The same factors also affect the inflammatory response, especially the inhibitory effect of salivary gland extracts on macrophage production of inflammatory cytokines (Bowman *et al.* 1997). Tick saliva also contains factors specific to inhibiting the inflammatory response. At least some Ixodid saliva contains inhibitors of neutrophils or a histamine antagonist (Ribeiro *et al.* 1990, Chinery and Ayitey-Smith 1977). Kininase activity in *Ixodes dammini* saliva has been found that deactivates bradykinin, a molecule involved in pain (Ribeiro 1987). Immunoglobin G binding proteins have been found in both the salivary glands and hemolymph of ticks and may be useful for capturing and excreting host IgG (Wang and Nuttall 1995). Other unique peptides have been

identified in tick saliva for which no definitive function has been determined (Bowman *et al.* 1997).

The pharmacologically active arsenal of compounds in tick saliva which an Ixodid tick uses to evade host detection and reaction creates an environment at the feeding lesion agreeable to certain pathogens that normally would be detected and destroyed by the host's defenses (Nuttal and Labuda 2004). Originally it was observed that uninfected Ixodid ticks feeding on animals inoculated with both salivary gland extracts and Thogoto virus acquired the virus 10 times more effectively than uninfected ticks feeding on animals inoculated only with the Thogoto virus (Jones *et al* 1989). Later, this phenomenon of enhanced transmission by salivary gland extracts was coined "saliva-activated transmission," or SAT (Nuttall and Jones 1991). Thus far, SAT has been implicated in the pathogenesis of at least one other virus and a number of bacteria for which ticks are vectors (Nuttall and Labuda 2004).

A Predominant ~90 kDa Immunoreactive Protein

Studies have reported the presence of a ~90 kDa protein in salivary gland extracts of *Rhipicephalus appendiculatus* that is immunoreactive to antiserum from rabbits on which the ticks fed (Venable *et al* 1986, Shapiro *et al* 1987, Jaworski *et al* 1992). It has been proposed that the 90 kDa antigen is a part of the attachment cement (Venable *et al* 1986, Shapiro *et al* 1987). Indeed, a 94 kDa protein was recognized by antiserum raised against solubilized cement (Shapiro *et al* 1989). However, the antigen that reacts with the 90 kDa protein does not react with a 94 kDa protein in an earlier study (Shapiro *et al* 1987, 1986). Shapiro and coworkers also mention that tick cement material may contain antigens that are not functional components of cement but ellicit an immune response (1986). Wang and Nuttall did not identify a 90 kDa protein in the

saliva of *Rhipicephalus appendiculatus*, but observed a 98 kDa protein dominant in saliva, salivary gland extracts, and hemolymph immunoreactive to serum from immune guinea pigs (1994).

A Newly Found Hemolymph Protein

A heme-lipoprotein (designated HeLp) has recently been isolated and partially characterized from the hemolymph of the hard tick *Boophilus microplus* (Maya-Monteiro *et al.* 2000). The protein, weighing an estimated 354 kDa (an average of the values given by native PAGE and gel filtration chromatography) and having a density of 1.28 g/ml, consists of two subunits with molecular weights of 103 and 92 kDa as revealed by SDS-PAGE and is the most abundant protein in the hemolymph based on concentration (Maya-Monteiro *et al* 2000). Both carbohydrates (3% (w/w)) and lipids (33%) compose HeLp, and 2 moles of heme associate with 1 mole of the native protein, although *in vitro* it may bind 6 additional moles (Maya-Monteiro *et al* 2000). Edman degradation was performed on both subunits, and unique N-terminal sequences of both subunits were obtained (Maya-Monteiro *et al* 2000).

Maya-Monteiro and coworkers suggest that HeLp is involved in heme transport and have demonstrated that heme bound to HeLp is transported to the ovaries during oogenesis (2000). The fact that at least one Ixodid tick cannot synthesize its own heme supports this (Braz *et al* 1999, Maya-Monteiro *et al* 2000). However, HeLp is the major protein component of male and unfed female hemolymph as well, suggesting HeLp has other roles (Maya-Monteiro *et al* 2000). In another study it was demonstrated that free heme promotes oxidation of phosphotidylcholine liposomes via the formation of free radical species significantly more than HeLp alone (with its "intrinsic" heme groups) (Maya-Monteiro *et al* 2004). When additional heme was added to

HeLp, the protein similarly reduced oxidation of the liposomes (when compared to equivalent concentrations of free heme) until the heme:HeLp ratio increased to 6:1 (Maya-Monteiro *et al* 2004). It was further demonstrated that the HeLp antioxidant character involved a lag phase in which no oxidation of liposomes occurred, suggesting the presence of an antioxidant particle associated with HeLp (Maya-Monteiro *et al* 2004). It was suggested that cholesterols may act as the antioxidant in HeLp (Maya-Monteiro *et al* 2004). In addition to its antioxidant role, HeLp in the presence of free heme has been demonstrated to prevent mammalian red blood cell lysis (and thus probably protects tick cells from lysis) that occurs in the presence of heme alone (Maya-Monteiro *et al* 2004). Maya-Monteiro and coworkers also demonstrated that HeLp prevents heme reactivity with lipid hydroperoxides that can result in free radical species and the destruction of heme itself (2004).

A protein very similar to HeLp was isolated and partially characterized from the hard tick *Dermacenter variabilis* and the soft tick *Ornithodoros parkeri* and hypothesized to be a carrier protein (designated DvCP and OpCP, respectively) (Gudderra *et al* 2001, Gudderra *et al* 2002a). Gudderra and coworkers found that DvCP is the most abundant plasma protein for both unfed and fed adult males and females, but only trace amounts were found in eggs (2001). The protein run on native gel electrophoresis appeared to be 210 kDa, but when run on a gel filtration column, the calculated molecular weight of the protein was 340 kDa (Gudderra *et al* 2001). No explanation is provided for this discrepancy (Gudderra *et al* 2001). OpCP appeared to be approximately 500 kDa on native-PAGE in one study (Gudderra *et al* 2001), but 668 kDa in another (Gudderra *et al* 2002a). Cp consists of two major subunits weighing 98 kDa and 92 kDa in *D. variabilis* and 114 kDa and 93 kDa in *O. parkeri* as revealed by SDS-PAGE, but Gudderra and coworkers mention a minor subunit weighing 48 kDa for OpCp (Gudderra *et al* 2001,

2002a). The density of DvCp was determined to be 1.25 g/ml (Gudderra *et al* 2001). It was determined that DvCp is a lipoprotein and contains carbohydrates, and that it binds heme as evident through absorption spectra and the brown coloring of the DvCp bands in unstained gels (Gudderra *et al* 2001). Edman degradation was performed on both DvCp subunits and N-terminal sequences were obtained, revealing that the lower subunit N-terminus has 61% identity to a fairy shrimp (*Streptocephalus proboscideus*) hemolymph biliprotein, artemocyanin (Gudderra *et al* 2001).

The N-termini of the DvCp subunits are nearly or completely identical to those of the *B. microplus* HeLp subunits (Maya-Monteiro *et al* 2000, Gudderra *et al* 2001, Gudderra *et al* 2002b, Dillwith and Madden, unpublished). This, along with the similar properties of the two proteins, suggests that HeLp and Cp are species-specific variants of the same protein (Gudderra *et al* 2002a, 2002b, Maya-Monteiro *et al* 2004, Madden *et al* 2002).

Maya-Monteiro *et al.* (2000) characterized the lipid content of *B. microplus* HeLp: 55% of the lipid content was phospholipid, while the remaining 44% was neutral lipid. Most of the neutral lipid (33% of the total lipid) was cholesterol ester, and cholesterol oleate was identified by mass spectrometry (Maya-Monteiro *et al.* 2000). The remaining lipids identified included free cholesterol, free fatty acid, and triacylglycerol (Maya-Monteiro *et al.* 2000). Gudderra and colleagues' characterization of DvCP lipids (2001) was somewhat different: triacylglyerol, cholesterol, and phospholipids were found, but not cholesterol oleate. Also, monoglycerides were found to associate with CP, although these lipids were not found in whole plasma.

Gudderra and group (2002a) have studied the distribution of DvCp and in fat body, salivary gland, ovary, and muscle tissue homogenates of partially fed and replete females of *D. variabilis*. In all cases the protein is more abundant in the plasma than in any tissue studied

(Gudderra *et al* 2002a). In partially fed females, DvCp was more abundant in the salivary glands than any other tissue studied (Gudderra *et al* 2002a). In replete females, DvCp is again more abundant in salivary gland tissue, but except for a faint band in the fat body homogenate lane, DvCp was not detected electrophoretically in any other tissue (Gudderra *et al* 2002a). Overall, DvCp is more abundant in the studied tissues of partially fed female than those of replete females (Gudderra *et al* 2002a). In the same study, the tissue distribution of OpCp in *O. parkeri* salivary gland, fat body, ovary, and coxal fluid of females 30 days after a blood meal and the OpCp in whole body homogenates of unfed and fed adult females and males was studied (Gudderra *et al* 2002a). Again, OpCp was found to be most abundant in the plasma (Gudderra *et al* 2002a). Except for the plasma, only OpCp from coxal fluid was electrophoretically resolved (Gudderra *et al* 2002a). Until a recent study (Madden *et al.* 2002), the presence of HeLp/Cp had not been studied in Ixodid tick saliva.

HeLp Precursor cDNA Cloned

Until recently, no full length sequence information was available for HeLp. Nene *et al.* (2004) published 2 EST consensus sequences from *Rhipicephalus appendiculatus* salivary gland transcript corresponding to a nearly complete HeLp precursor sequence. Later, a full-length *D. variabilis* HeLp (DvCp) precursor sequence (GenBank Acession No. DQ422963) was cloned. Apparently, HeLp is translated as a single, continuous polypeptide with the HeLp-B sequence preceding that of HeLp-A. Preliminary sequence analysis reveals that HeLp is a vitellogenin homologue. Seehuus *et al.* (2006) have recently demonstrated that vitellogenin in honeybees behaves as an antioxidant, supporting the hypothesis of Maya-Monteiro *et al.* (2004) that HeLp serves a similar role in ticks.

The ~90 kDa Salivary Protein May Be HeLp

A recent study investigates the predominant 95 kDa protein in *Amblyomma americanum* and *A. maculatum* tick saliva (when run on an 8-16% gradient SDS-PAGE gel) (Madden *et al* 2002). Each protein resolves into 2 distinct bands corresponding to 92 and 98 kDa on a 6% polyacrylamide gel (Madden *et al* 2002). Madden *et al.* suggest that the doublet corresponds to recently characterized HeLp (Cp) of the hemolymph and the 90 kDa so-called cement protein of tick saliva (Madden *et al* 2002, Madden *et al* unpublished). Indeed, the N-terminal sequences of the *A. americanum* and *A. maculatum* doublets are identical and nearly identical to those published for *B. microplus* HeLp and *D. variabilis* Cp (Madden *et al* unpublished).

Further evidence suggests that the HeLp in Ixodid tick saliva and hemolymph may be identical to the 90 kDa protein recognized by immune rabbit serum. Like HeLp, the 90 kDa saliva protein is glycosylated (Jaworski 1991). In an immunoblotted SDS-PAGE gel containing *D. variabilis* salivary gland extract, antiserum raised against *R. appendiculatus* recognizes a doublet (Jaworski *et al* 1992). Also, serum obtained from rabbits immunized with the 90 kDa protein reacts with a 300 kDa protein that Shapiro recognizes as a potential precursor for the 90 kDa protein (Shapiro *et al* 1987, Jaworski 1991). This may be non-denatured HeLp. Shapiro and coworkers (1987) and others (Jaworski *et al* 1992) note the importance of the 90 kDa antigen in inducing host immune response. Such proteins would be abundant in tick saliva and vital to tick feeding. HeLp qualifies as such as it has been characterized as the most abundant protein in Ixodid tick hemolymph and saliva and is an important heme carrier (Maya-Monteiro *et al* 2000, Gudderra *et al* 2001, Maya-Monteiro *et al* 2004).

Summary

The importance of understanding tick feeding is underscored by the economic and medical impact caused by these parasites and vectors of diseases. Crucial to tick feeding and survival are the compounds within the salivary excretions which allow a tick to even feed for weeks without significant host reaction and facilitate pathogen proliferation. Of these secretions a 90 kDa protein is predominant. Evidence suggests that this protein may be recently characterized HeLp or CP. HeLp binds heme and lipids and is glycosylated. Because of its abundance and highly conserved nature, and because Ixodid ticks cannot synthesize their own heme, this protein is probably vital to tick feeding and thus a significant protein to study for the ultimate goal of tick control.

CHAPTER 3

RESEARCH METHODS AND MATERIALS

Objective 1. Compare the structure of HeLp from six Ixodid species common to the USA.

Experimental Plan. HeLp has been reported to be the most abundant hemolymph protein in *Boophilus microplus* and *Dermacentor variabilis* (Maya-Monteiro *et al* 2000, Gudderra *et al.* 2001). Madden *et al.* (unpublished) have found that HeLp is the most abundant protein in both the hemolymph and saliva of *Amblyomma americanum*. Its abundance in species studied, its iron and lipid binding properties, and its very presence in the saliva are all evidence that HeLp is a protein crucial to tick survival. The conservation of such a protein among tick species is hypothesized. The purpose of this study is to examine the hemolymph and saliva of six US Ixodid tick species (*Amblyomma americanum*, *A. maculatum*, *Dermacentor andersoni*, *D. variabilis*, *Ixodes scapularis*, and *Rhipicephalus sanguineus*) for the presence of HeLp using SDS-PAGE. The structure of the HeLp subunits from these species was examined using Edman microsequencing and MALDI-TOF mass spectrometry.

Tick Rearing

Ticks were reared on rabbits and sheep at the Oklahoma State University Central Tick Rearing Facility using the methods described by Patrick and Hair (1975). Larvae were reared on mixed-breed rabbits and nymphs and adults on ewes. Partially engorged adult female ticks were used for hemolymph and saliva collection.

Tick Hemolymph Collection

Tick hemolymph was collected as described by Gudderra *et al* (2001). The leg was cut at the coxal-trochanteral joint and pressure was applied to express the hemolymph that was collected into a microcapillary tube. The hemolymph was diluted into an equal volume of Phosphate Buffered Saline, pH 7.4 (Sigma), and stored at -70 °C.

Tick Saliva Collection

Following hemolymph collection, saliva was collected from partially fed adult females as described by Ribeiro *et al.* (1992). Salivation was induced by injecting ticks immobilized on tape with 20 µl of a solution of 1 mM dopamine, 1 mM theophylline, and 1% dimethyl sulfoxide in PBS (Needham and Sauer 1979). Afterwards, each tick was injected with 10 µl of the above solution at 10 minute intervals for up to one hour. Colored saliva was discarded, and the rest was pooled into a microfuge tube that was frozen at -70 °C.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described by Laemmli (1970). For good separation of the HeLp subunits, tick hemolymph was electrophoretically separated on a resolving gel composed of 6.5% acrylamide in 375 mM Tris·HCl (pH 8.8) with a stacking gel consisting of 4.0% acrylamide in 125 mM Tris (pH 6.8). Prior to loading, samples were added to at least 5µl of a 4X strength sample buffer (28.8 mM 2-mercaptoethanol, 125mM Tris·HCl (pH 6.8), 20% (v/v) glycerol, and 6% SDS) and heated to 95°C. The reducing agent (2-mercaptoethanol) was excluded in some instances to resolve the proteins under non-reducing conditions. In instances when proteins were resolved under both

reducing and non-reducing conditions, lanes containing non-reduced samples were at least two lanes away (i.e., one lane is skipped) from lanes containing reduced sample to prevent the leaching over of reducing agent.

Gels were poured and run using a BioRad® Mini-PROTEAN® II electrophoresis system. The buffer chambers were filled with cold running buffer (25 mM Tris-HCl, 192 mM glycine, 0.1% SDS (w/v)) and the gels run at a constant 180 V for approximately one hour. Gels were stained overnight in a 0.1% Coomassie blue R-250 (Wilson 1983), 40% methanol, 10% acetic acid solution; destained 1 to 3 hours in a 40% methanol, 10% acetic acid solution; and stored in a 5% acetic acid solution at 4 °C.

N-terminal Sequencing of Protein Bands

For sequencing of proteins by Edman degradation, gels underwent a pre-run for 1 hour at 50 volts at 10°C with a pre-run buffer consisting of 375 mM Tris-HCl, 0.1% (w/v) SDS, 0.1 mM Na-thioglycolate solution at pH 8.8. Following pre-run, pre-run buffer was removed and replaced with normal running buffer and the samples were loaded. SDS-PAGE of proteins was achieved as described above. Protein bands from gels were electroblotted onto a PVDF membrane from an SDS-PAGE gel using the BioRad® Mini-PROTEAN® II transfer cell as described by Wilkins *et al.* (1996) and Dunn (1999), except that transfer occurred at 15 V overnight at 10°C. Membranes were stained with Coomassie R-250, and the bands of interest were excised and submitted to the Oklahoma State University Recombinant DNA/Protein Resource Facility for micro-Edman sequencing.

In Situ Digests

In-gel tryptic digests of proteins on Coomassie-stained gels were performed according to Jensen *et al.* (1999), and Shevchenko *et al.* (1996). All microfuge tubes and pipette tips used for trypsin digests were washed with 100 µl of 0.1% trifluoroacetic acid in a 50% acetonitrile solution. HeLp subunit bands were excised and washed according to Jensen *et al.* (1999), except that the gel slices were washed in a 50% acetonitrile, 25 mM ammonium bicarbonate solution. The proteins were reduced and alkylated *in situ* and the gels were dried according to Jensen *et al.* (1999).

The gel pieces were reswelled and the proteins digested according to Shevchenko *et al.* (1996). The proteins were digested in 40 µl of 0.025 ug/ul porcine trypsin (in 25 mM ammonium bicarbonate and Promega® trypsin buffer) overnight at 37 °C. The supernatant and the following extracts were pooled in a microfuge tube. The peptide fragments were extracted twice with 50 µl of 0.1% TFA, 50% acetonitrile solution for 30 minutes each and once more with 50 µl of acetonitrile for 20 minutes. The pooled supernatants were dried down with nitrogen gas to a volume of 2 to 5 µl.

***In Silico* Digests**

In silico tryptic digests of conceptual protein sequences were performed with the Protein Prospector MS Digest (of the UCSF Mass Spectrometry Facility) program at http://prospector.ucsf.edu/ucsfhtml4.0/msdigest.htm. Virtual digest parameters will allow for one missed cleavage, acrylamide modified cysteines, oxidation of methionine, and phosphorylation of serine, threonine, and tyrosine. The N-termini of the peptides was assumed to be hydrogen and the C-termini free acids. Peptides were assumed to be in the MH+ charge state.

MALDI-TOF Mass Spectrometry

Extracted protein digests were spotted on an Applied Biosystems® 96 well X2 plate compatible with an Applied Biosystems Voyager DE™-PRO Biospectometry™ Workstation. Recrystallized α-cyano-4-hydroxycinnamic acid (Beavis *et al.* 1992) was used for matrix. The matrix solution was prepared by dissolving 5 mg of matrix into 1 ml of a 0.1% TFA, 50% acetonitrile solution. An external calibrant was prepared by mixing 1 µl of calibrant peptides (containing des-Arg1-Bradykinin 1, angiotensin, fibrinopeptide, ACTH 1-17, and ACTH 18-39 peptides) obtained from Sigma with 50 µl matrix solution. Extracted digests (0.5 µl) were spotted on even rows of a MALDI plate and calibrants (1.0 µl) were spotted in close proximity on the odd rows. The digests were allowed to dry and 0.5 µl of matrix solution was spotted over the digest spots. An Applied Biosystems Voyager DE™-PRO Biospectometry™ Workstation was used in reflection mode for MALDI-TOF analysis. Data Explorer, the program accompanying the Voyager instrument for data analysis, was used to obtain monoisotopic *m/z* values. Traces were subjected to the following processing: the baseline was corrected for a peakwidth of 32, flexibility of 0.5, and a degree of 0.1. The traces were subjected to noise filtering/smoothing using the default settings. Traces were deisotoped, and the approximately 50 most intense peaks above the background peaks were selected for monoisotopic peak lists.

MALDI-TOF Mass Spectrometry Data Analysis

Monoisotopic peak lists were generated for each protein digest and from each list peaks corresponding to "contaminants" (i.e. trypsin, matrix, and keratin) were removed. For comparison of both subunits of HeLp between species, at least 3 digests were performed for each

subunit of each species. The lists for a protein subunit in each species were aligned according to *m/z* values. An *m/z* value in one list was considered to represent the same peptide fragment in another list of the same subunit and species if the standard deviation of all the *m/z* values for a candidate single peptide was less than 100 ppm, in accordance with the mass spectrometer's machine specifications. Using these criteria, master peak lists consisting of mean monoisotopic *m/z* values representing specific peptides were generated for each subunit of each species. *m/z* values were removed from a master list if the values did not occur in at least 50% of the lists (e.g. 2 out of 3 or 3 out of 4 lists). Peak master lists for a particular subunit were compared between species. Also, traces of the peptide mass fingerprints were visually inspected for similarities between species for a particular subunit.

Objective 2. To examine the composition of native HeLp in Amblyomma americanum.

Experimental Plan. At the commencement of this project, it was known that HeLp, a heme glycolipoprotein, binds lipids and heme iron and consists of two subunits of about 100 and 90 kDa (Maya-Monteiro *et al* 2000, Gudderra *et al* 2001, Gudderra *et al* 2002). Otherwise, very little was known about the structure of HeLp. The following experiments seek to clarify the composition of the native protein, especially in *A. americanum*.

First, hemolymph and salivary proteins were resolved under non-reducing conditions with SDS-PAGE to examine HeLp subunit association with regards to intermolecular disulfide linkages. In addition, a partial reduction of *A. americanum* HeLp-B using a serial dilution of reducing agent was performed to determine the number of HeLp-B subunits composing a single HeLp molecule.

Preliminary data showed other HeLp-related peptides in tick hemolymph and saliva. Therefore, these proteins were examined in reduced and unreduced tick hemolymph and saliva using SDS-PAGE, Western blotting, MALDI-TOF MS, and Edman microsequencing. These proteins may be clues as to how HeLp is processed.

Hemolymph and salivary proteins separated by SDS-PAGE were stained with a glycoprotein staining kit. Data concerning the glycosylation of the HeLp subunits were used to interpret anomalous molecular weight measurements.

Density gradient ultracentrifugation was used to isolate *A. americanum* HeLp and to determine its density in both hemolymph and saliva.

Tick Rearing

Ticks were reared on rabbits and sheep at the Oklahoma State University Central Tick Rearing Facility as described above.

Tick Hemolymph and Saliva Collection

Tick hemolymph and saliva was collected as described.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis and Staining

SDS-PAGE and staining with Coomassie blue R-250 was performed as described above. Silver staining of gels was accomplished according to Blum *et al.* (1988). After fixing, gels were washed in 50% ethanol twice for 20 minutes. Following development, Gels were rinsed in water and development stopped in 1.5% EDTA for 2-3 minutes. Gels were then rinsed with water and stored in a 5% acetic acid solution. Glycosylation of proteins on an SDS-PAGE gel was detected

using Pro-Q® Emerald 300 Glycoprotein Gel and Blot Stain Kit from Invitrogen. Protocol included with the kit was followed.

To study high molecular weight, non-reduced precursors of HeLp, bands excised from gels containing non-reduced proteins were incubated with a 4 X loading buffer with 2-mercaptoethanol for 10 minutes, sliced into smaller pieces, loaded into wells of a new gel, and overlayed with new loading buffer. SDS PAGE was then performed as described above.

Western Analysis

Western blots were performed as described by Towbin and Gordan (1984). For Western analysis, proteins were transferred to nitrocellulose membranes as described by Towbin *et al.* (1979) using the BioRad® Mini-PROTEAN® II transfer cell at a constant 10 V. Primary antibodies were from rabbit serum obtained by submitting our HeLp II N-terminal sequence (FEVGKEYVYKYKGTL) to Sigma Genosys (http://www.sigmaaldrich.com/Brands/Sigma_Genosys.html) for peptide synthesis and antibody production. The secondary anitibodies, goat anti-rabbit peroxidase conjugate obtained from Sigma, were visualized with 1-chloro-4-napthol substrate (Nakane 1968).

N-terminal Sequencing of Protein Bands

The N-terminal sequences of proteins were obtained as described above.

In Situ Digests

In-gel tryptic digests of proteins on Coomassie-stained gels were performed as described above.

MALDI-TOF Mass Spectrometry

MALDI-TOF MS analyses of digested proteins was performed as described above.

Density Gradient Ultracentrifugation

Centrifuged hemolymph and saliva proteins were separated via a KBr density gradient achieved in an ultracentrifuge. Prior to centrifugation, saliva was dialyzed against 1X stacking buffer and concentrated by ultrafiltration. A KBr solution made from KBr and saline (11.42g/L NaCl, 100 mg/L Na_2EDTA) that has a density of 1.18 g/ml was carefully layered over a KBr solution (in the centrifuge tube) with a density of 1.30 g/ml made from KBr and hemolymph or saliva. Beckman® OptiSeal™ Polyallomer 4.9 ml centrifuge tubes were used. The tube was carefully placed in a Beckman® VTi65.2 rotor and balanced with a tube containing similarly layered solutions. Samples were centrifuged in a Beckman® L8-70M ultracentrifuge for 90 minutes at 60,000 rpm and 8°C. 200 µl fractions were collected and analyzed with a Carl Zeiss Jena refractometer and a Hewlett Packard 8453 spectrophotometer.

Objective 3. To obtain nucleotide sequence information for *A. americanum* HeLp precursor transcript.

Experimental Plan. Obtaining cDNA sequence information for a protein provides an amino acid sequence, information which is fundamental for understanding secondary and tertiary structure and even function. A combination of the *D. variabilis* HeLp precursor cDNA sequence, the *A. americanum* HeLp tryptic peptide mass fingerprints, and the N-terminal sequences of HeLp subunits made cloning of *A. americanum* HeLp possible without library screening.

Degenerative primers based on sequences shared by the *D. variabilis* and *A. americanum* proteins as seen in comparisons of peptide mass fingerprints and from subunit N-terminal sequences were designed. Also, a stretch of sequence of *A. americanum* HeLp precursor cDNA (GenBank accession no. CX765178) obtained from a local EST database was used to design a gene-specific primer. Completion of HeLp cloning was obtained using 3'- and 5'-RACE.

Tick Rearing

Ticks were reared on rabbits and sheep at the Oklahoma State University Central Tick Rearing Facility as described above.

Salivary Gland Dissections

Partially fed *A. americanum* adult females were dissected according to Aljamali *et* al (2002) to obtain salivary glands. The dissection buffer was composed of 100 mM morpholinopropane sulfonic acid buffer (MOPS), pH 6.8, with 20 mM ethylene-bis-oxyethylene-nitrilotetraacetic acid (EGTA). Each group of 10 whole salivary glands was placed in microfuge tubes containing 300 µl of RNA*later*® (Ambion) and stored at -20°C for later use.

RNA Extraction

An RNAqueous Kit (Ambion) was used to isolate total RNA from partially-fed adult female salivary glands. Approximately 14 salivary glands were homogenized in 200 µl of lysis/binding buffer. Instructions included with the kit were followed. RNA, if not used immediately, was stored at -20°C.

RNA Gel Electrophoresis

The quality and quantity of RNA isolated was estimated with denaturing gel electrophoresis. A 10X RNA running buffer (200 mM MOPS, 10 mM EDTA, 50 mM sodium acetate, pH 7.0 with NaOH) was prepared for RNA gel electrophoresis reagents. 1% agarose (w/v) gels (125 ml) were made with nuclease free UltraPure™ agarose (Invitrogen), 12.5 ml of 10X RNA running buffer, and 108 ml nuclease free water. After boiling and cooling to 60 °C, to the agarose solution 6 ml of 35% formaldehyde was added. A sample master mix consisting of 50 µl 10 RNA running buffer, 250 µl deionized formamide, 90 µl formaldehyde, 108 µl nuclease free water, and 6µl of ethidium bromide solution (10 mg/ml) was used for denaturing RNA samples. Approximately 5 µg of RNA sample was added to 14 µl of RNA master mix and the mixture was heated to 65°C for five minutes. After heating, 4 µl of RNA loading dye (0.25% bromophenol blue, 0.25% xylene cyanol, and 20% ficoll, made with DEPC-treated water) was added to the sample mixture, which was then loaded into a gel well. An Invitrogen RNA ladder was loaded onto the gel for RNA size reference. A BIO-RAD Mini Sub Cell and components were used for electrophoresis. Gels were run immersed in a cold, circulating 1X RNA running buffer at 70V for approximately 1.5 hours. Gels were viewed with UV light.

cDNA Synthesis

Single-stranded cDNA was made from total RNA obtained from partially fed female *A. americanum* salivary glands using SuperScript™ III (Invitrogen) reverse transcriptase and oligo$(dT)_{20}$ (Invitrogen). Invitrogen's included protocol was followed, and cDNA was stored at minus 20°C.

Polymerase Chain Reaction

A GC-RICH PCR system (Roche-Applied Sciences) was used for polymerase chain reaction (PCR) reagents and the included protocols were followed. Typically, no GC-RICH resolution solution was needed for PCR amplification. Reactions occurred in a PTC™-100 Programmable Thermal Controller (MJ Research) with a heated lid. Cycling conditions suggested in the PCR kit protocol were followed.

PCR Clean-up

PCR products were purified using a ChargeSwitch® PCR clean-up kit (Invitrogen). The included protocol was followed.

DNA Gel Electrophoresis

Separation of PCR products or restriction digest fragments was achieved with agarose gel electrophoresis. 1.5% agarose (w/v) gels were made with 0.75 g nuclease free UltraPure™ agarose (Invitrogen), 0.75 g MetaPhor™ agarose (Midwest Scientific), 10 ml 10X TBE running buffer (90 mM tris, 80 mM boric acid, 3 mM EDTA, pH 8.35), and 90 ml sterile water. After boiling and cooling the gels to 60°C, 3 µl ethidium bromide solution (10 mg/ml) was added to the gel mixture before pouring. Gels were cooled at 4°C for 30 minutes before running. Approximately 100 ng or less of DNA sample was loaded per lane in the gel. DNA samples were mixed with 1 µl Blue/Orange 6X loading dye (Promega) and diluted with 1X TBE running buffer if necessary. 8 µl or more of the sample mixture was loaded per lane. A 1 kb Plus DNA Ladder (Invitrogen) and a Low DNA Mass Ladder (Invitrogen) were used for DNA size and quantity estimation, respectively. A BIO-RAD Mini Sub Cell and components were used for

electrophoresis. Gels were run immersed in a cold 1X TBE running buffer at 120V for approximately 1 hour. Gels were viewed with UV light.

Gel Extraction of Nucleic Acid Bands

DNA bands of interest that needed to be purified were excised from agarose gels and stored at -20°C until needed. A QIAquick® gel extraction kit (Qiagen) was used to isolate DNA from gel pieces, following the provided protocol.

Cloning of PCR Products

A GeneJET™ PCR cloning kit (Fermentas) was used for cloning PCR products and the included protocols were followed. PCR products were assumed to have sticky ends; therefore the sticky-end protocol was followed with the GeneJET™ kit. In one instance, a TOPO TA Cloning® kit (Invitrogen) was used to clone a PCR product following the included instructions. Ligated plasmids were purified via two phenol:chloroform:isoamyl alcohol (25:24:1) extractions followed by an ethanol precipitation. TOP10 electrocomp™ (Invitrogen) *E. coli* cells were transformed with plasmids with a BioRad Gene Pulser® electroporator in 2 mm Gene Pulser® cuvettes. The electroporator was set to 2.5V, a capacitance of 25 Faradays, and resistance of 200 Ω. Transformation and culturing practices described in the TOPO TA cloning manual were followed. Bacteria were plated on agarose made with Luria Broth containing ampicillin, restreaked after an overnight incubation at 37°C, and, following another overnight incubation, grown in LB ampicillin broth overnight at 37°C.

Plasmid Purification

For isolation of plasmid clones, a Wizard® *Plus* Minipreps DNA Purification System (Promega) was used and the included instructions followed. Prior to plasmid purification, an aliquot of bacteria suspended in LB ampicillin broth was mixed with glycerol for a final concentration of 15% glycerol and stored at -80°C. The remaining suspension was used for plasmid purification.

Restriction Digest Analysis

BamHI and EcoRI restriction enzymes and kits from Promega were used to analyze plasmids for the presence of and quantify inserts. Included protocols were followed, and the MULTI-CORE™ buffer was used as the enzymes (1:1) were used in combination.

DNA Sequencing

PCR products or purified plasmids were submitted to the Oklahoma State University Recombinant DNA/Protein Resource Facility for sequencing. Internal primers or the pJet1 Forward or pJet1 Reverse sequencing primers included with the GeneJet™ PCR cloning kit were used with an Applied Biosystems 3730 DNA Analyzer.

Cloning Strategy

Degenerative primers J26 (5’ AAG AGG ATG GCN GCN CTN TGG GC 3’) and L35 (5’ TCC TCG ATC TCY TTN CGY TCY TT 3’) designed from the internal sequences (KRMAALWA and KERKEIEDSL, respectively) obtained from MALDI-TOF MS and the *D. variabilis* HeLp cDNA sequence or Edman degradation sequencing were used to clone an initial

internal cDNA fragment (DC or "Dillwith Clone"). Sequence information obtained from clone DC, along with sequence information from the N-terminal end of the HeLp precursor or the 3' end sequence residing in GenBank (GenBank accession no. CX765178, nicknamed VJ) was used to design primers for cloning two more stretches of cDNA: stretch E13DCR, upstream of DC, was amplified with primers E13 (5' TTC GAR GTI GGI AAR GAR TAY GT 3'), designed from the N terminal sequence FEVGKEYV, and DCR (5' TCG CTG TCG TTG TGG AAG ACG 3'), obtained from sequence DC; while stretch DCFVJR, downstream of DC, was amplified with primers DCF (5' GCT CGC ACG ACA GCT ACT TCC 3'), obtained from sequence DC, and VJR (5' CCA TCG ACG CGA ACA ATC AGG 3'), obtained from sequence VJ.

Rapid Amplification of 5' or 3' Ends (5' or 3' RACE)

The 5' and 3' ends of the HeLp precursor transcript were obtained using Clontech's SMART™ RACE cDNA amplification kit, and the included protocols were followed. For amplification of the 5' and 3' ends of the cDNA, gene specific primers GSP-5' (5'-GGC TGA GTG GGG GTG CGT CTC CTG GGCG-3') and GSP-3' (5'-CCG TGC CAC CGG CAA CCC TGC GCT CGC C-3') were used, respectively. PCR reagents, thermocycler programs, and cloning were utilized as described above.

Sequence Analysis and Contig Construction

Sequence chromatograms received from the OSU Recombinant DNA/Protein Resource Facility were uploaded into Invitrogen's contig assembly program ContigExpress and then edited and assembled into a contig with the software.

CHAPTER 4

RESULTS AND DISCUSSION

Objective 1

A survey of HeLp in American Ixodid ticks

SDS-PAGE analysis of the hemolymph of six Ixodid species revealed that proteins with molecular weights similar to those of the HeLp subunits reported for *Boophilus microplus* (Maya-Montiero *et al.* 2000), *Dermacentor variabilis* (Gudderra *et al.* 2001), and *Amblyomma americanum* (Madden *et al.* 2006) are the most abundant proteins in the hemolymph of all species examined (Fig. 1). The most aberrant molecular weights of these HeLp subunits occur in the hemolymph of *A. maculatum*, where both subunits of HeLp migrate as one intensely staining band, and *Ixodes scapularis*, the smaller HeLp subunit of which has an exceptionally lower molecular weight. For better subunit separation in further studies, *A. maculatum* HeLp subunits were separated with SDS-PAGE under non-reducing conditions. Electrophoresis of Ixodid tick saliva results in stained gels similar to those seen in the hemolymph study (data not shown).

Hemolymph and salivary proteins separated by SDS-PAGE were electroblotted onto a PVDF membrane, excised, and submitted for Edman microsequencing. The N-terminal ends of both HeLp subunits are nearly identical for the respective subunits in all species examined and are in agreement with the sequences published for *Boophilus microplus*, *R. appendiculatus*, and *D. variabilis* (Table 1 and Table 2).

Protein bands were excised from SDS-PAGE hemolymph gels and subjected to *in situ* digestions with porcine trypsin. Digested proteins were extracted and analyzed with a MALDI-TOF mass spectrometer, and peptide mass fingerprint monoisotopic peak lists were generated for the most intense peaks. Also, the partial available sequence of *R. appendiculatus* HeLp-B and the full length sequences of *R. appendiculatus* HeLp-A and *D. variabilis* HeLp-A and HeLp-B were subjected to *in silico* tryptic digestions with "virtual" conditions similar to the those of the real digestions. The peak lists from real and virtual digests were compared and compiled into common peptide mass lists (Tables 3 and 4). Our results suggest that many peptide masses, and thus peptide sequences, are shared between these species for each subunit, and in cases in which real digestion peak masses are matched with virtual masses, actual sequence data and even information on phosphorylation is provided for these species. In some instances there are ambiguities in matching real peak masses with virtual masses. These result from the coincidence of different peptides, real or virtual, having very similar masses. This is unavoidable because the mass spectrometer has limited mass accuracy (100 ppm using external standards). However, the matched peptide masses are so numerous that it would be safe to assume that most are correctly matched. In the case of *D. variabilis* when both a real peptide mass fingerprint and a virtual mass fingerprint of a known sequence are available, it is evident that not all real peptide masses can be matched with virtual masses. This may result from allelic variations in HeLp genes, post-translational modifications of the proteins, and unequal or favored peptide ionization and flight in the mass spectrometer.

The N-terminal sequences and the peptide mass fingerprints of HeLp-A and HeLp-B reveal that these subunits are highly conserved between Ixodid species. The conservation of HeLp and its abundance in both Ixodid hemolymph and saliva underscores its importance in

ticks and its status as a promising vaccine target. Predictably, the HeLp-A and HeLp-B subunits had the most shared peptides between species in the same genus (i.e., *Amblyomma spp.* And *Dermacentor spp.*); *D. andersoni* and *D. variabilis* HeLp-A or HeLp-B have nearly identical peptide mass fingerprint traces when inspected visually (Fig. 2). The *I. scapularis* HeLp subunits shared the least number of peptides with the real and virtual digests of HeLp-A and HeLp-B from the other species, in addition to having the most unique N-terminal sequences for both subunits and a lower HeLp-B molecular weight. Our results seem logical in light of *I. scapularis* being the only member of the Protstriata used in this study.

Ultimately the HeLp precursor cDNAs will be cloned for each of these species, and while such an accomplishment would be important for understanding HeLp and for species comparisons, mass spectrometry will remain a valuable tool for determining aspects of the structure of this protein and of others that cannot otherwise be seen from a nucleotide sequence. Meanwhile, mass spectrometry may facilitate the cloning of these precursors by providing actual protein sequence data for the design of primers to amplify the cDNA or even the gene.

Objective 2

A. americanum *HeLp-Related Peptides*

Preliminary Western blots of saliva and hemolymph polyacrylamide gels led us to the investigation for HeLp-related peptides in these sources. SDS-PAGE analysis of hemolymph and salivary proteins from partially fed *A. americanum* was conducted under reducing conditions to visualize the less abundant proteins (Fig. 3a). Rabbit polyclonal antibodies raised against the N-terminal peptide sequence of HeLp-B were available for Western analysis, revealing a number of immunoreactive bands (Fig. 3b). In both hemolymph and saliva, the HeLp-B band and a 210 kDa

band were visible on Western blots. In the saliva a 300 kDa band, and in the hemolymph an ~80 kDa band were also detected. Bands on duplicate gels corresponding to the immunoreactive bands were excised for MALDI-TOF analysis. Up to six bands were excised from a lane: HeLp-A and HeLp-B as controls, the 300 kDa and 210 kDa immunoreactive bands, and two bands (75 and 80 kDa) ambiguously associated with the ~80 kDa immunoreactive band (Fig. 3).

MALDI-TOF mass spectrometry analysis was applied to the tryptic digests of these proteins, revealing that all are related to HeLp (Table 5). The resulting fingerprints of proteins other than HeLp-A or HeLp-B appeared to be a "mix" of the HeLp-A and HeLp-B fingerprints—that is, the 300, 210, 80, and 75 kDa proteins all contain tryptic digest peptides common to both HeLp-A and HeLp-B (Table 5). No protein digest resulted in an appreciable number of peptides not found in a typical HeLp-A or HeLp-B digest. Analysis of the N-terminal sequences of the 210 kDa and 300 kDa proteins reveals that the 210 kDa protein has an N-terminus identical to that of HeLp-B and the 300 kDa species has two N-termini identical to the HeLp-A and HeLp-B N-terminal sequences (data not shown). Similarly, Maya-Monteiro *et al.* (2000) observed a higher molecular weight band in purified HeLp samples separated with SDS-PAGE containing mixed HeLp-A and HeLp-B N-termini.

That a 210 kDa peptide has the N-terminus of HeLp-B, contains HeLp-A and HeLp-B, and weighs approximately as much as the sum of the weights of HeLp-A and HeLp-B suggests, along with evidence from the HeLp precursor cDNA sequences of *R. appendiculatus* and *D. variabilis*, that both subunits originate from a common precursor peptide. The presence of a 300 kDa peptide in reduced *A. americanum* hemolymph and saliva with digest peptides correlating only to HeLp-A and HeLp-B complicates our picture: evidence from the cDNA sequences suggests that a HeLp precursor could not be more (or very much more when taking post-

translational modifications into account) than 210 kDa. However, as Edman microsequencing reveals, a mixture of the HeLp-A and HeLp-B N-terminal sequences occurs for this protein, suggesting it may result from the attachment of a 210 kDa precursor to a HeLp-A molecule with a non-reducible kind of crosslinkage. It is unknown whether or not this form (or the 210 kDa form) of HeLp has a function. However, a crayfish plasma clotting protein homologous to vitellogenin (like HeLp) forms intermolecular crosslinkages via a transglutaminase-mediated clotting reaction (Hall *et al.* 1999). It is tempting to entertain the possibility that HeLp may also "clot", resulting in the macroscopic cement cones observed around tick mouthparts and in the feeding lesion. The 80 and 75 kDa peptides may be degradation products of HeLp, as they contain parts of both subunits but are smaller than either one. Interestingly, Jaworski *et al.* (1992) report that antibodies raised against a 90 kDa cement antigen react with a 70 kDa protein from cement cones. Whether or not HeLp is related to tick cement protein needs to be investigated.

HeLp structural characterization

Thus far all electrophoresis studies of HeLp have been performed in native conditions or in reducing, denaturing conditions. To further understand HeLp structure, we have studied *A. americanum* hemolymph and saliva with denaturing SDS-PAGE under non-reducing conditions (Figure 4). In non-reducing conditions, an abundant protein migrated at a position corresponding to 80 kDa and another abundant protein at 275 kDa. MALDI-TOF mass spectrometry (Table 3) and Western blotting (data not shown) confirmed that the 80 kDa band is HeLp-A alone and that the 275 kDa band is composed entirely of HeLp-B. Apparently, the HeLp-A subunit does not crosslink with other subunits via disulfide bridge formation, but rather, contains internal linkages due to it migrating as if it were 80 kDa under non-reducing conditions. Indeed, the HeLp

precursor sequence reveals that HeLp-A contains a von Willibrand factor type D domain which contains a number of internal disulfide bridges.

A distinct, high molecular weight triplet (Fig. 4a) (designated in the order from smallest to greatest mobility as NR-1, NR-2, NR-3) immunoreactive to polyclonal anti-N-terminal HeLp-B antibodies occurred above the 275 kDa band in a Western blot (not shown). No molecular weight estimate is offered for these three proteins as their relative mobilities occur well outside the standard curve. MALDI-TOF mass spectrometry revealed that all three proteins are composed of both HeLp-A and HeLp-B (Table 5). Similar banding patterns resulted with the other species examined except that more than 3 higher molecular weight bands occurred (it is difficult to count these bands which are grouped closely together) (Fig. 4b). The composition of the high molecular weight triplet of non-reduced *A. americanum* hemolymph was examined by excising the bands and reducing them (Fig. 5). Silver staining of the reduced proteins separated by SDS-PAGE revealed that the highest molecular weight member of the triplet, NR-1, is composed mostly of a 210 kDa peptide. This peptide has the same mobility of the 210 kDa protein composed of HeLp-B and HeLp-A run on the same gel. The products of the reduction of NR-2 are a 300 kDa peptide and a 93 kDa peptide. The 300 kDa peptide has the same mobility as the reduced 300 kDa peptide containing HeLp-A and HeLp-B, and the 93 kDa peptide migrates to the same position as reduced HeLp-B. The products of the reduction of NR-3 have apparent molecular weights of 210 kDa, corresponding to the same peptide in NR-1, and 93 kDa, migrating at the same position as HeLp-B. When the reduced 210 and 300 kDa bands (both of which contain HeLp-A and HeLp-B) were exposed to the same reducing conditions as members of the non-reduced triplet and are examined by SDS-PAGE again, most of the samples migrate

once again to positions corresponding to 210 and 300 kDa respectively, although a fainter band appeared for both samples at a position that may correspond to HeLp-A.

The presence of a high molecular weight triplet in non-reduced hemolymph and saliva samples separated by SDS-PAGE, each member of the triplet containing HeLp-A and HeLp-B and composed of 210 or 300 kDa peptide, is puzzling. The presence of the precursors (if they are indeed precursors) in both the hemolymph and the saliva suggests that either the proteins are processed in the hemolymph and/or saliva, or that they are merely the result of "leakage" of unprocessed intracellular HeLp and perhaps have no real function. The fact that the HeLp precursors of *R. appendiculatus* and *D. variabilis* both contain a basic processing site (not shown) common to substrates of subtilisn-like proprotein convertases in the Golgi complex supports the latter argument (Barr 1991, Rouille *et al.* 1995).

Non-reduced SDS-PAGE reveals that the HeLp-B subunits may possibility form trimers (275 kDa ≈ 3 × 93 kDa). It seems logical that a trimer of HeLp-B and a single HeLp-A subunit could result in the approximately 354 kDa native HeLp protein that Maya-Monteiro *et al.* (2000) isolated from *Boophilus microplus* (especially if the molecular weights derived from the migrations of the non-reduced subunits were used [275 kDa + 80 kDa = 355 kDa]). However, this does not take the lipid moieties that Maya-Monteiro *et al.* (2000) quantifies for *B. microplus* HeLp and Gudderra *et al.* (2001) mentions for *D. variabilis* into account. Also, HeLp-A and HeLp-B appear to occur in a stoichiometry of 1:1 (Maya-Monteiro *et al.* 2000). Certainly it is important to note that methods used thus far for obtaining the molecular weights for native HeLp and its subunits (e.g., gel filtration chromatography, pore-limiting PAGE, and SDS-PAGE) do not necessarily reveal true molecular weights. In order to clarify the quaternary association of the HeLp-B subunits, a partial reduction of the unreduced HeLp-B subunit complex excised from a

polyacrylamide gel was performed (Fig. 6). A number of HeLp-B protein bands excised from a gel were reduced (in a dilution series of DTT) and alkylated (to prevent reoxidation of cysteines) *in situ*, placed in the wells of a new polyacrylamide gel, and analyzed again by SDS-PAGE. Our results reveal that, although different oxidation states of HeLp-B were "created," no major intermediate existed between the 275 and 90 kDa forms of HeLp-B—that is, no supposed dimer intermediate occurred between the alleged HeLp-B trimer and HeLp-B monomer. These results seem to suggest that HeLp-B occurs as a cross-linked dimer as opposed to a trimer. A dark band and a lower trailing smear occurred in lanes in which HeLp-B was partially reduced. These regions are probably artifacts resulting from HeLp-B subunits being partially reduced and reforming new, biologically insignificant intermolecular disulfide linkages with numerous HeLp-B molecules (i.e., the subunits became "sticky" and "polymerized").

The staining of *A. americanum* hemolymph and salivary proteins on a polyacrylamide gel with a glycoprotein staining kit (Fig. 7b) or Schiff's reagent alone (not shown) reveals that HeLp-B is glycosylated, HeLp-A staining substantially lighter (Fig. 7b). However, the calculated molecular weight of unmodified *R. appendiculatus* or *D. variabilis* HeLp-A derived from the precursor (92.4 kDa or 92.3 kDa, respectively) is smaller than the observed molecular weights of *R. sanguineus* or *D. variabilis* HeLp-A with SDS-PAGE (102 kDa and 100 kDa, respectively), suggesting that it, too, may be glycosylated. The NetNGlyc 1.0 Server of the Center for Biological Sequence Analysis (http://www.cbs.dtu.dk/services/NetNGlyc/) (Gupta *et al.* 2004) predicts 3 N-glycosylation sites (residues 241, 480, and 558) on both *R. appendiculatus* and *D. variabilis* HeLp-B proteins. However, an unglycosylated peptide (mass 1026.5849) corresponding to residues 235 to 242 occurs in *A. maculatum* HeLp-B, suggesting a glycosylation does not occur here for this species. This peptide was not found in other real

digests. Only one N-glycosylation is predicted for HeLp-B at residue 777 for *D. variabilis* or 778 for *R. appendiculatus*. That HeLp-B is substantially glycosylated may explain its unusually high apparent molecular weight in unreduced form as observed with SDS-PAGE.

Density Gradient Ultracentrifugation

At this time there is no mention of the presence of HeLp in Ixodid saliva in the literature. As HeLp has been proposed to be a carrier protein, it would not be unreasonable to hypothesize that HeLp may act to traffic lipids in between the feeding lesion and the tick hemolymph and tissues. To investigate the density of HeLp in saliva as well as hemolymph, saliva and hemolymph were fractionated in a KBr density gradient achieved by density gradient ultracentrifugation (Figs. 8 and 9). For hemolymph, the relative abundance of HeLp in a fraction was determined by its absorbance at 280 nm (for total protein) and 400 nm (for heme iron). From the density gradient a density of 1.270 g/ml was calculated for HeLp. Five fractions containing the highest titers of HeLp were separated with SDS-PAGE (Fig. 10), revealing a banding pattern not unlike crude hemolymph. Very faint bands corresponding to about 300 kDa and 210 kDa were present, as well as a number of darker staining lower molecular weight bands. These smaller proteins, which appear to be present in crude hemolymph and saliva as well, might either be the products of HeLp degradation that occurred after centrifugation, or may have intrinsic densities similar to that of HeLp so that they copurify with HeLp. The 200 and 300 kDa species, though crosslinked or incompletely processed HeLp, might have a similar tertiary structure to native HeLp and thus the lipid-binding capabilities of the latter.

Fractionation of HeLp in saliva proved much more difficult. For one, though HeLp is the most abundant protein in tick saliva, saliva (at least that which is collected by our methods) is

orders of magnitude more dilute than hemolymph. Second, a species absorbing substantially at our wavelengths of interest interfered with our studies. Although dialysis significantly reduced the quantity of this absorbing species, this further diluted our saliva. Absorbance traces of fractionated saliva revealed maxima at 275 nm and 395 nm, probably corresponding to HeLp. A density of 1.274 g/ml could be derived for salivary HeLp, a value similar to that of HeLp in hemolymph. The described experiments need to be repeated and the lipids analyzed in HeLp fractions to determine if the lipid compositions of HeLp are similar in both sources.

Objective 3

HeLp Precursor Transcript Cloning and Analysis

Fragments of the *A. americanum* HeLp transcript were amplified by PCR using the knowledge gained from the MALDI-TOF mass spectrometry studies, Edman degradation microsequening, and *D. variabilis* HeLp transcript sequence information. Fragments were cloned and sequenced, and a contig presumed to contain all of the HeLp precursor cDNA was constructed (Fig. 11). At the time of the construction of the contig, the aforementioned VJ sequence (GenBank accession no. CX765178) was used as the sequence of the 3' end. This EST contains a long string of A residues, and so the exact length of the poly-A tail is uncertain. As expected, 25 bases upstream of the poly-A tail begins a poly-A signal sequence (AATAAA). An alignment of the *A. americanum* HeLp (AaHeLp) cDNA with the *D. variabilis* HeLp (DvHeLp) cDNA reveals that the nucleotide sequences have a 75.2% identity (not shown). In the alignment, DvHeLp has one 3 residue-long gap in the 5' untranslated region, AaHeLp a 3 residue gap in the

HeLp-A coding region, and both have a number of small gaps in the 3' untranslated regions of the sequences.

Conceptual translations of the AaHeLp and DvHeLp precursor sequences were aligned with only one gap resulting: at amino acid residue 1093, DvHeLp has an aspartic acid for an additional residue, resulting in a protein that is one residue longer than that of *A. americanum* (not shown). The precursor proteins share 75.1% sequence identity, similar to the sequence identity shared between the cDNAs. If similar amino acids are included in the identity figure, 85.7% of the residues are conserved. Clearly the two proteins are nearly identical.

BLAST searches using either the DvHeLp or AaHeLp precursor conceptual translations reveal that HeLp is homologous to vertebrate, arthropod and nematode vitellogenin (data not shown). A Conserved Domain Search (Marchler-Bauer and Bryant 2004) of AaHeLp, DvHeLp, and their vitellogenin homologues reveal that the proteins share very similar domain architectures (Fig. 12). A multiple alignment of a number of vitellogenin-related N-terminal domains with those of AaHeLp and DvHeLp was performed (data not shown). To illustrate the sequence similarity and perhaps the phylogenetic relationships between these proteins, a tree was constructed using the neighbor-joining method (Fig. 13). The unrooted tree depicts three major groups: the vertebrate and nematode vitellogenins; the insect vitellogenins; and the arachnid vitellogenins, as well as the arachnid, crustacean, and even vertebrate vitellogenin-related proteins. The latter group contains the human microsomal triglyceride particle transfer protein (MTP), the entire protein of which only shares its N-terminal domain with the vitellogenins. According to this tree, although related to vitellogenin, the HeLp ancestors did not split from the tick vitellogenin clade recently. Another tree (not depicted) constructed with the same parameters but from an alignment from the entire lengths of the proteins (excluding MTP)

groups tick vitellogenin with the insect vitellogenins, but in a way that is deeply rooted. In this scenario, the HeLp ancestor split even before the rise of the tick vitellogenins. The bootstrap values of the second tree are generally poorer, but otherwise the trees are very similar.

The relatedness of vitellogenin and HeLp may give insight into the processing of HeLp. In insects, vitellogenins usually contain a dibasic site (RXXR or RXRR) that is cleaved by proprotein subtilisn-like convertases in the Golgi body (Barr 1991, Rouille *et al.* 1995, Chen *et al.* 1997). As mentioned in objective 1 for *D. variabilis* and *R. appendiculatus*, *A. americanum* precursor HeLp contains the sequence RFPR preceding the N-terminus of the mature A subunit (Fig. 11 and 14). Convertase cleavage sites are usually flanked by charged, serine-rich regions forming β-turns and occur near or on one of these β-turns (Rholam *et al*. 1986; Brakch *et al.* 1993). In AaHeLp precursor, numerous serines and charged residues indeed flank this cleavage site (Fig. 11).

HeLp Modeling

With HeLp precursor sequences available, it becomes possible to model the tertiary structure of HeLp, provided a structure is available for a homologous protein or a domain found within HeLp. At this time, the only HeLp homologue with a high resolution structure solved is *Ichthyomyzon unicuspis* lipovitellin (Protein Data Bank ID LSH1) (Fig. 15). This lipovitellin is an incomplete representative of the vitellogenin structure because a number of disordered regions are found on the protein chain and lipovitellin itself is a fully processed vitellogenin with stretches of sequence cleaved from the chain (Anderson *et al.* 1998). Lamprey lipovitellin consists of 4 characterized folds that do not necessarily reside on the same protein chain (Anderson *et al.* 1998) (Fig. 15): The N-sheet consists mainly of beta strands forming a tight

barrel. From the N-sheet arises a helical domain containing at least 18 parallel helices fanning in a near-semicircle out from one end of the N-sheet. At the C-terminus of the helical domain arises a small C-sheet, which folds over part of the inner surface of the helical domain, forming an alpha-beta sandwich. The large A-sheet forms a half beta barrel with a large cavity. From the N-terminal portion of the A-sheet, a strand-turn-strand is inserted into a cavity of the N-sheet, forming what Anderson *et al.* describe as a ball and socket (1998). The entire structure has the appearance of a wrist (N-sheet) with an open hand bent out (helical domain and C-sheet) holding the A-sheet (Fig. 15).

Though a structure of a vitellogenin is available for HeLp modeling, it is necessary that both proteins are similar enough so that such a model would be reliable. An alignment of the HeLp precursor with *Ichthyomyzon unicuspis* vitellogenin precursor (Fig. 16) reveals that the proteins align well, though a number of gaps are present. Generally, the region of HeLp that corresponds to HeLp-B is longer than the same region of lipovitellin, although the insertions occur in regions not interrupting beta strands or alpha helices in lipovitellin (Fig. 17). As a result, the structure of HeLp-B may not be significantly altered from the template structure of lipovitellin. For the region of lipovitellin that aligns with the HeLp-A region, HeLp is much shorter, and the missing regions correspond to numerous beta strands of the A-sheet in lipovitellin (Fig. 17). This problem, in combination with the fact that much of the second half of lipovitellin is not present in the structure, suggests that this vitellogenin would not model HeLp-A well. However, the alignment does give credibility to a model of HeLp-B.

HeLp-B was successfully modeled using the CPH 2.0 homology modeler (Lund *et al.* 2002) at http://www.cbs.dtu.dk/services/CPHmodels/ (Fig. 18). The model, which contains structural elements homologous to lipovitellin's N-sheet, helical domain, and C-sheet, includes

nearly all of the HeLp-B residues. This model is an excellent tool for interpreting data produced from the second objective. The model places the 9 cysteines of HeLp-B in positions that are reasonable (Fig. 19)—most pair with another cysteine within a distance that allows disulfide linkages. C-428 and C-443, as well as C-456 and C-470, pair in such a manner in buried regions of the helical domain (Figs. 20 and 21, respectively). C-145 and C-170 pair well within the N-sheet, with only C-170 having the possibility of rotating the sulfur group into the solvent accessible region of the subunit (Fig. 22). C-186, C-189, and C-616 congregate in close proximity (Fig. 23). C-616 resides on the alpha helical domain, while the other two lie about a loop in the N sheet. C-189 and C-186 (Fig. 24A), as well as C-616 and C-186 (Fig. 24B), can adopt configurations allowing a disulfide bridge, but C-616 and C-189 are too far apart, at least in the rigid model, to form such a linkage (Fig 24). The sulfur of C-189 is buried, and so could not engage in an intersubunit disulfide linkage, unlike C-186 or C-616 (Fig 24). Anderson *et al.* (1998) report that lipovitellin dimers make contact mainly at the N sheet and alpha helical domains. For the HeLp-B model, C-170, C-186, and C-616 are in positions near the bent junction of the N sheet and alpha helical domain that would correspond to lipovitellin's dimerization surface. For C-186 to participate in an intersubunit disulfide linkage, all intramolecular disulfide linkages would be sacrificed in the C-186-189-616 region (Fig 24). Both C-186 and C-616 (Fig. 25) are in a pocket; therefore any linkage of either of those residues with another HeLp-B subunit would require flexing of the HeLp-B subunits or a modification of the model. If C-170 sacrificed its linkage with C-145, it would be in a relatively easily accessible configuration for intersubunit disulfide bridging (Fig. 26). It is even possible that a number of these cysteines form linkages with those of another subunit. Regardless of the scenario, at least one of these cysteines mentioned (C-170, 186, or 616) is responsible for the HeLp-B intersubunit

linkage known to exist. Site-directional mutagenesis studies might clarify the subunit associations.

The HeLp-B model can be a used to test the credibility of the glycosylation sites predicted in objective 2. As is the case for DvHeLp-B, the NetNGlyc 1.0 Server of the Center for Biological Sequence Analysis (http://www.cbs.dtu.dk/services/NetNGlyc/) (Gupta *et al.* 2004) predicts 3 possible glycosylation sites (residues 241, 480, and 558) for the AaHeLp-B sequence (Fig. 27). The server gives N-241 the highest score and 9 out of 9 neural networks concur with the prediction. N-480 has a lower score, but 8 of 9 neural networks support the prediction. Residue N-558 has the lowest score, and only 4 of 9 networks agree with the prediction. Figure 28 depicts the position of the first predicted site of glycosylation. Residue N-241 is located on the surface of the N-sheet opposite the predicted surface of dimerization, and glycosylation here would not appear to interfere with HeLp-A subunit association with HeLp-B, if indeed HeLp-A associates with HeLp-B as the homologous regions do in lamprey lipovitellin. Residue N-480 is solvent accessible in the model (Fig. 29), and a glycosylation at this site would occur on the side of the helical domain opposite of the hypothesized HeLp-A binding site. It is not clear if a glycosylation here would interfere with, or even aide in, HeLp-B dimerization. Residue N-558 occurs on the same surface of the helical region as N-480, but closer to the possible intersubunit disulfide linkages (Fig. 30). However, this residue is near a “fingertip” of the helical domain, and so glycosylation here may not interfere with dimerization. None of the predicted sites of glycosylation have been disqualified using the HeLp-B model.

Lipid-binding in lipovitellin is an unresolved topic—the latest crystallographic data reveal only one layer and a few lipids from a second layer of hydrocarbons and phospholipids binding on the surface of the hydrophobic walls of the lipid-binding cavity composed by the A-

and C-sheets (Thompson and Banaszak 2002). Dimerization of lipovitellin closes one of the three entrances to the lipid binding cavity, resulting in two entrances accessible to the solvent in each subunit (Anderson *et al.* 1998). NMR-spectroscopy data have suggested that the phospholipids may occur as a monolayer or bilayer within the lipid-binding cavity (Banaszak and Seelig 1982). Likewise, the location and orientation of other lipids can only be speculated (Anderson *et al.* 1998). The lack of information on vitellogenin lipid-binding, along with the lack of structural information on the HeLp-A subunit that is hypothesized here to bind lipid and perhaps heme, leaves little for further understanding of HeLp lipid-binding. However, it is known that vitellogenins bind only up to 20% lipid (w/w) (16% in lamprey lipovitellin). Though the structure of native HeLp is unresolved, its homology to vitellogenin, along with molecular weight and density data obtained from this project, seem to suggest that the 33% lipid (w/w) value attributed to HeLp by Maya-Monteiro *et al.* (2000) is incorrect.

CHAPTER 5

CONCLUSIONS

Key to the survival of ticks and the vector status of the Ixodidae is tick saliva. It contains a poorly understood blend of biological molecules that allow exploitation of the host and silent infection by a wide spectrum of pathogens. Despite the variety of compounds found in the saliva, the secretion is dominated by a large lipoprotein, heme lipoprotein. Amazingly, this protein is also the most abundant tick hemolymph protein. Its mere abundance, along with its lipid-binding and heme-binding character, suggest that it is indispensable and worthy of study. In this study this protein has been surveyed in six important American ticks and has been further studied in detail in the lonestar tick, *Amblyomma americanum*.

The survey of American ticks has revealed that it is similarly abundant in the saliva and hemolymph of all species studied, and is most likely highly conserved in sequence. The presence of this protein in a tick may in fact be a morphological prerequisite for belonging to the family Ixodidae. Its conservation and conserved abundance makes it amenable for study and qualifies it as a target for tick control.

This study has revealed that HeLp occurs as a heterodimer of homodimers (or vice versa), having the form $\alpha_2\beta_2$. The HeLp-B subunit weighs about 90 kDa and is moderately to heavily glycosylated. It is linked to another HeLp-B subunit via at least one disulfide bridge. The HeLp-A subunit weighs about 100 kDa and associates with HeLp-B noncovalently. It may be

glycosylated, and has a significant number of internal disulfide linkages, characteristic of the von Willebrand factor type D domain it contains.

A combination of Western blotting and MALDI-TOF mass spectrometry has revealed that HeLp-related peptides occur in hemolymph and saliva. In particular, a 210 kDa protein in hemolymph and saliva with the N-terminus of HeLp-B and peptides belonging to both HeLp-A and HeLp-B led to the hypothesis by our group that an A and B subunit occur as a single peptide precursor, even before the release of *D. variabilis* cDNA sequence. In a non-reduced denaturing gel, it was shown that these alleged precursors are linked with 300 kDa peptides to form different higher molecular weight species. Interestingly, the larger peptide has two N-termini: that of HeLp-B and that of HeLp-A. Furthermore, the presumed linkages within the 300 kDa species cannot be broken by reduction. These higher molecular weight HeLp species may be some form of unprocessed HeLp that has leaked from the processing machinery. Because HeLp is glycosylated, excreted, and probably a substrate of a proprotein convertase, it must be processed, at least partially, in the endoplasmic reticulum and Golgi body of the cell. Therefore, unprocessed HeLp that is excreted could not be further processed. Smaller molecular weight HeLp-related peptides occur in hemolymph and saliva and fractions containing HeLp purified by density gradient ultracentrifugation. These may be degradation products of HeLp.

The density gradient ultracentrifugation studies have shown that HeLp has a similar density, and thus, possibly a similar lipid load in the hemolymph and saliva. This does not mean that HeLp does not unload its lipid or heme, however. In both cases (and certainly for the saliva) the protein may not have arrived at its destination and gained or lost its cargo.

A cDNA sequence for *A. americanum* HeLp shows that it is homologous to vitellogenin, although it has not arisen recently from tick vitellogenin. Rather, its ancestor may have been present along with the tick vitellogenin ancestor before the rise of the tick clade. HeLp is related to vitellogenin not only by structure but by function: both are lipid carriers and bind heme (at least in ticks). The homology of vitellogenin and HeLp has allowed the modeling of the B subunit of the latter. The HeLp-B model has the appearance of an open palm. This palm, composed of the N-sheet (the "wrist"), a helical domain (the "fingers"), and the C-sheet (the "thumb"), probably "holds" the HeLp-A subunit, which most likely composes a lipid binding cavity as in vitellogenin. This model has supported the hypothesis that the B subunit composes crosslinked dimers and has supported the possible glycosylation sites.

The present study has given us a foothold onto a previously mysterious protein (though newer and more questions have arisen). It is a launch pad for other studies: site of synthesis studies (using the cDNA sequence for real-time PCR and Northern blotting); functional studies (using the cDNA sequence for RNA interference); site-directed mutagenesis studies for structural and functional information (e.g., testing cysteines required for HeLp-B subunit cross-linking or for removing glycosylations); lipid binding studies (using density-gradient ultracentrifugation and other techniques); HeLp-ligand and binding partner studies (using immunoprecipitation and affinity chromatography); gene cloning; elucidation of transcriptional regulation; development of a universal tick antigen for tick vaccines; and endless others leading in numerous directions. The understanding of this protein appears to be crucial to our understanding of Ixodid ticks and their transmission of pathogens.

LITERATURE CITED

Aljamali, M., Bowman, AS., Dillwith, JW., Tucker, JS., Yates, GW., Essenberg, RC., Sauer, JR., 2002. Identity and synthesis of prostaglandins in the lone star tick, *Amblyomma americanum* (L.), as assessed by radio-immunoassay and gas chromatography/mass spectrometry. Insect Biochem. Molec. Biol. 32, 331-341.

Altschul, SF., Gish, W., Miller, W., Myers, EW., Lipman, DJ., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403-10.

Anderson, TA., Levitt, DG., and Banaszak, LJ. 1998. The structural basis of lipid interactions in lipovitellin, a soluble lipoprotein. Structure. 6, 895-909.

Banaszak, LJ. and Seelig, J. 1982. Lipid domains in the crystalline lipovitellin/phosvitin complex: a phosphorus-31 and deuterium nuclear magnetic resonance study. Biochemistry. 21, 2436-43.

Barr, PJ., 1991. Mammalian subtilisins: the long-sought dibasic processing endoproteases. Cell. 66, 1-3.

Beavis, RC., Chaudhary, T., Chait, BT., 1992. α-cyano-4-hydroxycinnamic acid as a matrix for matrix-assisted laser desorption mass spectrometry. Org. Mass Spectrom. 27, 156–8.

Blum, H., Beier, H., Gross, HJ., 1988. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis. 8, 93-9.

Bowman, AS., Dillwith, JW., and Sauer, JR., 1996. Tick salivary prostaglandins: presence, origin and significance. Parasitol. Today. 12, 388-396.

Bowman, AS., Coons, LB., Needham, GR., Sauer, JR., 1997. Tick saliva: recent advances and implications for vector competence. Med. and Vet. Entomol. 11, 277-85.

Brakch, N., Rholam, M., Boussetta, H., and Cohen, P.1993. Role of b-turn in proteolytic processing of peptide hormone precursors at dibasic sites. Biochemistry. 32, 4925–30.

Braz, GRC., Coelho, HSL., Masuda, H., Oliveira, PL., 1999. A missing metabolic pathway in the cattle tick *Boophilus microplus*. Current Biology. 9, 703-6.

Chen, J., Sappington, TW., and Raikhel, AS. 1997. Extensive sequence conservation among insect, nematode, and vertebrate vitellogenins reveals ancient common ancestry. J. Mol. Evolution. 44, 440-51.

Childs, JE., Paddock, CD., 2003. The ascendancy of *Amblyomma americanum* as a vector of pathogens affecting humans in the United States. Ann. Rev. of Entomol. 48, 307-337.

Chinery, WA., Ayitey-Smith, E., 1977. Histamine blocking agent in the salivary gland homogenate of the tick *Rhipicephalus sangineus*. Nature. 265, 366-367.

Dunn, MJ. 1999. Electroblotting of proteins from 2-D polyacrylamide gels. In: Link, A.J. (Ed), 2-D proteome analysis protocols: Methods in Molecular Biology. Humana Press, Totowa, pp. 313-18.

Gale, A., and Ringdahl, E. 2001. Tick-borne diseases. Am. Fam. Physician. 64: 461-6, 468.

Gudderra, NP., Neese, PA., Sonenshine, DE., Apperson, CS., Roe, RM., 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.

Gudderra, NP., Sonenshine, DE., Apperson, CS., Roe, RM., 2002a. Tissue distribution and characterization of predominant hemolymph carrier proteins from *Dermacentor variabilis* and *Ornithodoros parkeri*. J. Insect Physiol. 48, 161-70.

Gudderra, NP., Sonenshine, DE., Apperson, CS., Roe, RM., 2002b. Hemolymph proteins in ticks. J. Insect Physiol. 48, 269-78.

Hackman, RH., Filshie, BK., 1982. The Tick Cuticle. In F. Obenchain and R. Galun (Eds.) Physiology of Ticks (p. 2). Pergamon Press.

Hall, M., Wang, R., van Antwerpen, R., Sottrup-Jensen, L., Soderhall, K., 1999. The crayfish plasma clotting protein: a vitellogenin-related protein responsible for clot formation in crustacean blood. Proc. Natl. Acad. Sci. USA. 96, 1965-70.

Higgins, D., Thompson, J., Gibson, T., Thompson, JD., Higgins, DG., Gibson, TJ., 1994. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-80.

Jaworski, DC., 1991. Molecular aspects of early-expressed Ixodid tick salivary gland antigens with emphasis on host response to tick feeding. Doctoral dissertation, Ohio State University. (p. 56).

Jaworski, DC., Rosell, R., Coons, LB., Needham, GR., 1992. Tick (Acari: Ixodidae) attachment cement and salivary gland cells contain similar immunoreactive polypeptides. J. Med. Ento. 29, 305-309.

Jensen, ON., Wilm, M., Shevchenko, A., Mann, M., 1999. Sample preparation methods for mass spectrometric peptide mapping directly from 2-D gels. In: Link, A.J. (Ed.), 2-D proteome analysis protocols. Humana Press, Totowa, pp. 519-520.

Jones, LD., Hodgson, E., Nuttall, PA., 1989. Enhancement of virus transmission by tick salivary glands. J. General Virol. 70, 1895-1898.

Jongejan, F. Uilenberg, G., 2004. The global importance of ticks. Parasitol. 129, S3-14. Cambridge University Press.

Karczewski, J., Waxman, L., Endris, RG., Connolly, TM., 1995. An inhibitor from the argasid tick *Ornithodoros moubata* of cell adhesion to collagen. Biochem. Biophys. Res. Comm. 208, 532-541.

Laemmli, UK., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 277, 680–85.

Lund, M., Nielson, C., and Lundegarrd, P.W. 2002. CPH models 2.0: X3M a computer program to extract 3D models. Abstract at the CASP5 conferenceA102, 2002.

Madden, RD., Cordill, WJ., Dillwith, JW. 2006. Occurrence of HeLp in the saliva of the lone star tick, *Amblyomma americanum*. Submitted.

Madden, RD., Sauer, JR., Dillwith, JW., 2002. A proteomics approach to characterizing tick salivary secretions. Exper. and Appl. Acarol. 28, 77-87.

Mans, BJ. and Neitz, AWH., 2004. Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. Insect Biochem. Molec. Biol. 34, 1-17.

Marchler-Bauer, A. and Bryant, S.H. 2004. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32(W), 327-331.

Maya-Monteiro, CM., Daffre, S., Logullo, C., Lara, FA., Alves, EW., Capurro, ML., Zingali, R., Almeida, IC., Oliveira, PL., 2000. HeLp, a heme lipoprotein from the hemolymph of the cattle tick, *Boophilus microplus*. J. Biol. Chem. 275, 36584-9.

Maya-Monteiro, CM., Alves, LR., Pinhal, N., Abdalla, DSP., Oliveiro, PL., 2004. HeLp, a heme-transporting lipoprotein with an antioxidant role. Insect Biochem. Mol. Biol. 34, 81-7.

Nakane, PK., 1968. Simultaneous localization of multiple tissue antigens using the peroxidase-labeled antibody method: A study on pituitary glands of the rat. J. Histochem. Cytochem. 16, 557.

Needham, GR., Saur, JR., 1979. Involvement of calcium and cyclic AMP in controlling Ixodid tick salivary fluid secretion. J. Parasitiol. 65, 531-42.

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 Biochem. Mol. Biol. 34, 1117-28.

Nuttall, PA. and Jones, LD., 1991. Mites and ticks as vectors of disease. In F. Dusbabek and V. Bukva (Eds.) Modern Acarology. Vol.2. (pp. 3-6). The Hague.

Nuttall, PA. and Labuda, M., 2004. Tick-host interactions: saliva-activated transmission. Parasitol. 129, S177-189

Obenchain, FD., and Galun, R., 1982. Preface. In: Obenchain, F., Galun, R. (Eds.), Physiology of Ticks. Pergamon Press, Elmsford, p. *vii*.

Packham, MA., 1994. Role of platelets in thrombosis and hemostasis. Canadian J. Pharmacol. 72, 278-284.

Patrick, CD., Hair, JA., 1975. Laboratory rearing procedures and equipment for multi-host ticks. J. Med. Entomol. 12, 389-90.

Radomski, MW., Palmer, RMJ., Moncada, S., 1987. The anti-aggregating properties of vascular endothelium: interactions between prostacyclin and nitric oxide. British J. Pharmacol. 92, 639-646.

Rholam, M., Nicolas, P., and Cohen, P. 1986. Precursors for peptide hormones share common secondary structures forming features at the proteolytic processing sites. FEBS Letters. 207, 1–6.

Ribeiro, JMC., Mather, TN., Piesman, J., Spielman, A., 1987. Dissemination and salivary delivery of Lyme disease spirochetes in vector ticks (Acari: Ixodidae). J.Med. Ento. 24, 201-205.

Ribeiro, JMC., 1987. Role of saliva in blood-feeding by arthropods. Ann. Rev. Ento. 32, 463-478.

Ribeiro, JMC., Evans, PM., McSwain, JL., Saur, JR., 1992. *Amblyomma americanum*: characterization of salivary prostaglandins E_2 and $F_2\alpha$ by RP-HPLC bioassay and gas chromatography mass spectrometry. Exp. Parasitol. 74: 112-16.

Rouillé, Y., Duguay, SJ., Lund, K., Furuta, M., Gong, Q., Lipkind, G., Oliva, AA., Jr., Chan, SJ., Steiner, DF., 1995. Proteolytic processing mechanisms in the biosynthesis of neuroendocrine peptides: the subtilisin-like proprotein convertases. Frontiers in Neuroendocrinol. 16, 322-61.

Sambrook, J., Fritsch, EF., Maniatis, T., 1989. Molecular Cloning. Cold Spring Harbor Laboratory Press.

Sappington, TW., Raikhel, AS., 1998. Molecular characteristics of insect vitellogenins and vitellogenin receptors. Insect Biochem. Mol. Biol. 28, 277-300.

Sauer, JR., McSwain, JL., Bowman, AS., Essenberg, RC., 1995. Tick salivary gland physiology. Ann. Rev. Ento. 40, 245-267.

Sauer, JR., Essenberg, RC., Bowman, AS., 2000. Salivary glands in Ixodid ticks: control and mechanism of secretion. J. Insect Phys. 46, 1069-1078.

Seehuus, S., Norberg, K., Gimsa, U., Krekling, T., and Amdam, G.V. 2006. Reproductive protein protects functionally sterile honey bee workers from oxidative stress. PNAS. 103: 962-7.

Shapiro, SZ., Voigt, WP., Fujisaki, K., 1986. Tick antigens recognized by serum from a guinea pig resistant to infestation with the tick *Rhipicephalus appendiculatus*. J. Parasitol. 72(3), 454-463.

Shapiro, SZ., Buscher, G., Dobbelaere, DA., 1987. Acquired resistance to *Rhipicephalus appendiculatus* (Acari: Ixodidae): identification of an antigen eliciting resistance in rabbits. J. Med. Entomol. 24, 147-154.

Shapiro, SZ., Voigt, WP., Ellis, JA., 1989. Acquired resistance to Ixodid ticks induced by tick cement antigen. Exper. Appl. Acarol. 7, 33-41.

Shevchenko, A., Wilm, M., Vorm, O., Mann, M., 1996. Mass spectrometric sequencing of proteins from silver-stained polyacryamide gels. Anal. Chem. 68, 850–8.

Sonenshine, DE., 1993. *Biology of Ticks*. Vol 2. Oxford University Press, New York, pp. 333-334.

Sonenshine, DE., 1991. Biology of Ticks. Vol 1. (pp. 3-4, 45, 51-66, 102-104). Oxford University Press.

Sonenshine, DE., 1993. Biology of Ticks. Vol 2. (pp. 333-334). Oxford University Press.

Thompson, JD., Gibson, TJ., Plewniak, F., Jeanmougin, F. Higgins, DG., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876-82.

Thompson, JR. and Banaszak LJ. 2002. Lipid-protein interactions in lipovitellin. Biochemistry. 41: 9398-409.

Towbin, H., Staehelin, T., Gordon, J., 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets : procedures and some applications. Proc. Natl. Acad. Sci., USA. 76, 4350.

Towbin, H., Gordan, J., 1984. Immunoblotting and dot immunoblotting: Current status and outlook. J. Immunol. Methods. 72, 313.

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

Venable, JH., Webster, P., Shapiro, SG., Voight, WP., 1986. An immunocytochemical marker for the complex granules of tick salivary glands which traces e-granule shedding to interstitial labyrinthine spaces. Tissue and Cell. 18, 765-781.

Walker, D.H. 1998. Tick-transmitted infectious diseases in the United States. Annu. Rev. Public Health. 19:237-69.

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

Wang, H., Nuttall, PA., 1995. Immunoglobin-G binding proteins in the Ixodid ticks, *Rhipicephalus appendiculatus*, *Amblyomma variegatum* and *Ixodes hexagonus*. Parasitol. 111, 161-165.

Whelan, AC., Richardson, LK., Wikel, SK., 1984. Ixodid tick antigens recognized by the infested host: immunoblotting studies. IRCS Med. Sci. 12, 910-911.

Wikel, SK., Ramachandra, RN., and Bergman, DK., 1994. Tick-induced modulation of the host immune response. International J. for Parasitol. 24, 59-66.

Wilkins, MR., Ou, K. Appel, RD., Sanchez, JC., Yan, JX., Golaz, O., Farnsworth, V., Cartier, P., Hochstrasser, DF., Williams, KL., Gooley, AA., 1996. Rapid protein identification using N-terminal “sequence tag” and amino acid analysis. Biochem. Biophys. Res. Comm. 221: 609-13.

Wilson, CM., 1983. Staining of proteins on gels: Comparison of dyes and procedures. Methods in Enzymol. 91, 236-47.

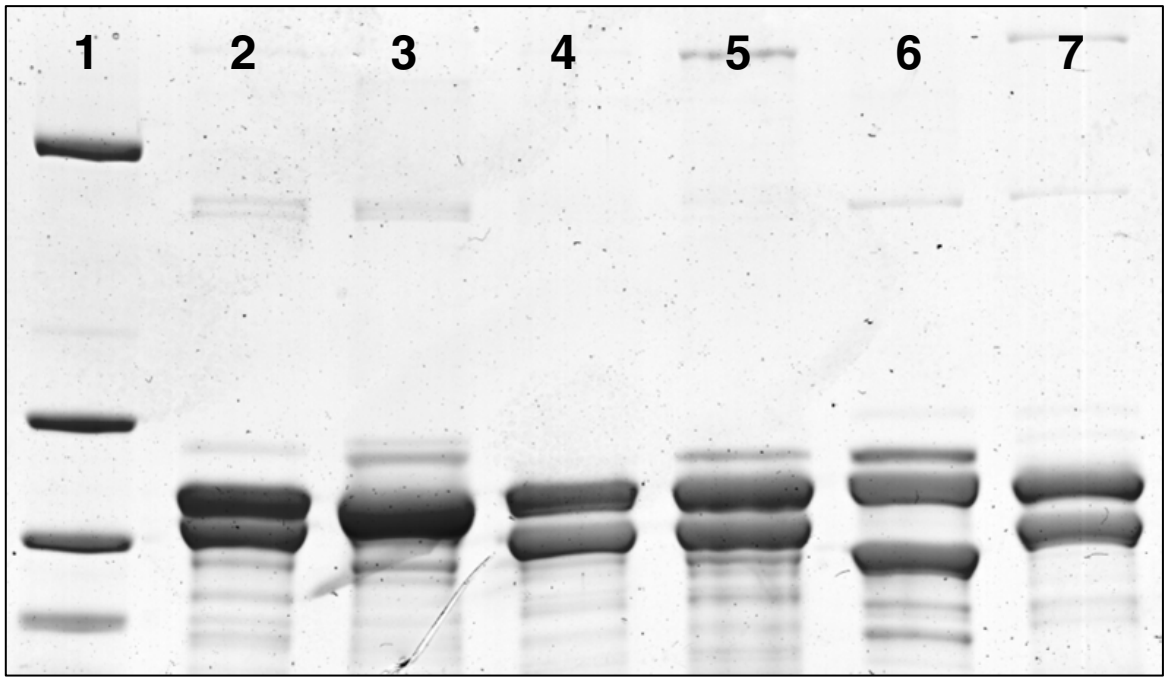


Fig. 1. 6.5% SDS-PAGE of the hemolymph of six Ixodid tick species (reducing conditions). Lane 1, Sigma® wide molecular weight standards (Catalogue no. M-4038), lane 2, *A. americanum*, lane 3, *A. maculatum*, lane 4, *D. andersoni*, lane 4, *D. variabilis*, lane 5, *I. scapularis*, lane 6, *R. sanguineus*.

Table 1: N-terminal sequences of HeLp-A in Hemolymph and Saliva

Species	Mass (kDa)	Source	N-terminal Sequence
A. americanum	99	Hemolymph	D A S A K E R K E I E D **S** L
A. americanum	99	Saliva	D A S A K E R K E I E D **S** L
A. maculatum	96	Hemolymph	D A S A K E R K E I E D
A .maculatum	96	Saliva	D A S A K E R K E I E D
D. andersoni	98	Hemolymph	D **T** S A K E R K E I E D A
D. andersoni	98	Saliva	D **X** S A K E R K E I E D A
D. variabilis	98	Hemolymph	D A S A K E R K E I E D
D. variabilis	98	Saliva	D A S A K E R K E I E D A
R. sanguineus	102	Hemolymph	D A S A K E R K E I E D
R. sanguineus	102	Saliva	D A S A K E R K E I E D
I. scapularis	100	Hemolymph	D A S A K **D L Q** E **V** E **A**
I. scapularis	100	Saliva	D A S A K **D L Q** E **V** E **A**
***B. microplus**	103	Hemolymph	D A S A K E R K E I E D A L
****D. variabilis**	98	Hemolymph	**X E P** A **N** E R **X** E I E D A L
*****R. appendiculatus**	?	Salivary Gland EST	D A S A K E R K E I E D A L

*From Maya-Monteiro *et al.* 2000.
**From Gudderra *et al.* 2001.
*** Consensus sequence obtained from BLAST results from The Institute of Genomic Research.

Table 2: N-terminal sequences of HeLp-B in Hemolymph and Saliva

Species	Mass (kDa)	Source	N-terminal Sequence
A. americanum	93	Hemolymph	F E V G K E Y V Y K Y K G T L
A. americanum	93	Saliva	F E V G K E Y V Y K Y K G T L
A. maculatum	94	Hemolymph	F E V G K E Y V Y K Y K
A. maculatum	94	Saliva	F E V G K E Y V Y K Y K
D. andersoni	91	Hemolymph	F E V G K E Y V Y K Y K
D. andersoni	91	Saliva	F E V G K E Y V Y K Y K
D. variabilis	92	Hemolymph	F E V G K E Y V Y K Y K
D. variabilis	92	Saliva	F E V G K E Y V Y K Y K
R. sanguineus	94	Hemolymph	F E V G K E Y V Y K Y K
R. sanguineus	94	Saliva	F E V G K E Y V Y K Y K
I. scapularis	88	Hemolymph	F E V G K **D** Y V Y **H** Y **N**
I. scapularis	88	Saliva	F E V G K **D** Y V Y **H** Y **N**
***B. microplus**	92	Hemolymph	F E V G K E Y V Y K Y K G T L
****D. variabilis**	92	Hemolymph	**V** E V G K E Y V Y **X** Y **X** G T L
*****R. appendiculatus**	?	Salivary Gland EST	F E V G K E Y V Y K Y K G T L

*From Maya-Monteiro *et al.* 2000.
**From Gudderra *et al.* 2001.
***Consensus sequence obtained from BLAST results from The Institute of Genomic Research.

Table 3: HeLp-A Common Monoisotopic Peptide Masses[a]

AVG	Std Dev (ppm)	Common	Ra Peak	Dv Peak	Translation	PO_4	start…end
927.4613		Is		927.5515	VTLAPLEGK		165…173
967.4945	22.92	Da, Dv		967.4638	HYLYESR		125…131
974.5357	11.26	Aa, Am					
981.5291		Rs	981.4794		HYLYETR		125…131
988.5168	40.49	Am, Dv, Is		988.4892 or 988.4951	LQFFYDR or DAPEKTAEK		368…374 or 54…62
989.5664		Rs	989.5784	989.5784	NYKPLLNK		510…517
1012.5129	17.22	Da, Dv, Rs	1012.4700 or 1012.4740	1012.4700 or 1012.4740	SSLGSGYDAR or EYDPVYAR		145…154 or 20…27
1020.5088	18.35	Da, Dv, Rs	1020.4638	1020.4638 or 1020.4719	FLEPDDER or MALQCNAGR		307…314 or 470…478
1024.5052	40.08	Aa, Am, Is					
1032.5598	19.14	Da, Dv		1032.5267	FPKPDWSR		395…402
1036.5314		Rs	1036.4892	1036.4668	FQFFYER or MALQCNAGR		368…374 or 470…478
1041.5300		Rs	1041.4906	1041.4618, 1041.4658, or 1041.5369	FPHPDWSR, STNNPSLTK, LTADVEWK, or NLATFYQGK	0,1,1,0	395…402, 657…665, 501…508, or 405…413
1094.5608	21.66	Dv, Rs					
1107.5950		Aa		1107.5482	HRHCDLQK		612…619
1111.5851	11.72	Dv, Rs	1111.5536	1111.5536	QAEFIYANR		97…105
1214.6890		Rs	1214.7122		FLRHSSRLAK		492…501
1347.7755		Am		1347.6657	EIEDALHIHDR		9…19
1361.7492		Rs	1361.6814		EIEDALHIHER		9…19
1421.7421	13.31	Aa, Da, Dv, Rs	1421.6602	1421.6602	YSFNHDLFNHK		357…367
1428.7696		Am		1428.6330	VKNLATFYQGK	2	403…413
1455.7137	18.29	Aa, Am, Da, Dv	1455.6612	1455.6361	LCGLCGDYNLDK or LCGLCGDYNLDR		754…765
1475.8624	9.58	Aa, Am		1475.7607	KEIEDALHIHDR		8…19
1600.8547		Rs	1600.7946		LQQMVGFAWTYTR		132…144
1653.9867	15.38	Aa, Da, Dv, Rs	1653.9427	1653.9427	LEILPVTEDSGLIVR		677…691
1700.9598		Da		1700.8566	LHLVSQVPWWDLK	1	549…561
1709.9237	42.74	Aa, Da, Dv, Is		1709.8916	LYYITINPHWHPR		270…282
1750.9010	11.90	Da, Dv		1750.8282 or 1750.8488	SNDQEGTKICLEASAK or **HWNVPTFSHMLEPR**		379…394 or 580…593
1766.9126	0.93	Da, Dv		1766.8437	HWNVPTFSHMLEPR		580…593
1805.0287		Am	1804.9063		YNGLHRMALKCQAGR		465…479
1847.0212		Am	1846.9717		YQALRPEHGGFFGVIR		525…540
1853.7578	20.41	Da, Dv					
1855.9991	11.31	Aa, Am		1855.8591	IDVTHGDNAEIHLDVK	1	108…123
1857.9550		Rs	1857.8547		AVDSWDFDESIFEAVK		36…51
1870.0178	26.75	Aa, Am, Da, Dv		1869.9876	YHALRPEHGGFFGIIR		524…539
1903.9629		Dv		1903.8367 or 1903.9125	AVDSWDFDESILDAIK or **GLHSGLDEFYHQMTLR**	1,0	36…51 or 252…267
1919.0937		Da		1918.9849	ICLEASAKFPKPDWSR		387…402
1919.9570		Dv		1919.9074	GLHSGLDEFYHQMTLR		252…267
1931.0805	11.53	Aa, Am					
1985.0130		Rs		1984.9356	ICLEASAKFPKPDWSR	1	387…402
2325.2672		Am		2325.0569	SSLGSGYDARTVISWPLDLK	2	145…164
2592.3384		Is		2592.1256	MALQCNAGREHGIPFNYYCMK		470…490
2598.4688		Am		2598.3397	SYGIYLTFNGNLLFIQTAPFYR		729…750

Table 4: HeLp-B Common Monoisotopic Peptide Masses[a]

AVG*	Std Dev PPM	Common	Ra Peak	Dv Peak	Translation	PO_4	start…end
901.5004	9.22	Da, Dv		901.4896	ALHYLER		511…517
906.5373		Am	906.4937		LEEFAIGK		93…100
913.5134		Aa	913.4856		SEAVKSHR		430…437
919.4719	54.26	Am, Rs	919.5075	919.5075	MAALWALK		531…538
944.5125	18.72	Da, Dv, Is		944.4954	NEQLWVR		101…107
974.4787	10.21	Da, Dv, Rs	974.4696	974.4696	SHDSYLPR		671…678
989.5116	6.41	Dv, Rs					
1008.5054	28.10	Aa, Am					
1011.5216	30.11	Aa, Am					
1026.5864		Am	1026.5849	1026.5849	QFHVILNR, QFHVLLNR		235…242
1037.5841		Is	1037.5281	1037.4991	WHANEPKR or NLWNFMGR		523…530 or 722…729
1053.5017	6.28	Da, Dv, Rs		1053.494	NLWNFMGR		722…729
1078.5344	26.04	Aa, Am					
1166.6627		Is	1166.5893		SLWNFMGRR		722…730
1301.6915		Am	1301.7404	1301.7404	MAALWALKQAAK		531…542
1314.6727	32.09	Aa, Am					
1331.8504	23.66	Aa, Dv, Is		1331.7939	GVLSLFQLDLVK		112…123
1345.8406		Rs	1345.8095		GILSLFQLDLVK		112…123
1387.7531	4.33	Da, Dv		1387.7374	YVLPLWETNPR		626…636
1397.7283		Is		1397.6037	SHREIWTSCK	1	435…444
1419.7535	9.78	Da, Dv	1419.7636		EYVYKYKGTLR		6…16
1423.6573	2.53	Aa, Am					
1430.6762	20.73	Da, Rs					
1440.6446	29.03	Aa, Am					
1457.7430		Aa	1457.8415		RMAALWALKQAAK		530…538
1471.7970	30.77	Aa, Am					
1499.7899		Aa		1499.8334	VFNRLVGPQPGSTK		708…721
1513.7918	17.33	Aa, Rs					
1529.7980	20.40	Aa, Rs	1529.7674		YDYGGLTLVEMIR		658…670
1551.8199		Is		1551.7854	YPCHKEIAQHLR		614…625
1558.8521		Am	1558.9321		GILSLFQLDLVKGR		112…125
1565.8356		Is		1565.8011	YPCHKEIAQHLR		614…625
1659.8858		Rs	1659.8264	1659.7661	EVSPALLQEIADSCK or VFNRLVGPQPGSTK	0,2	415…429 or 708…721
1673.9011		Rs	1673.842		EVSPALLQEIADSCK		415…429
1701.0247		Am	1700.9349		HPNRLVAHFGLTPNK		287…301
1766.9922		Is		1766.9302	SAGNLGTHKALHYLER		502…517
1847.9869		Dv		1847.9656	LEEFSIGKNEQLWVR		93…107
1849.8817	2.43	Da, Dv		1849.8608	DVEHEPEYESTVFIR		487…501
1963.0714		Da		1962.9326	SIALPVFHNTSEPSEVR	1	550…566
2106.1478	9.14	Da, Dv		2106.1096	GTLHVANPEQPLQASGIAFR		13…32
2186.1622		Aa		2186.0759	GTLHVANPEQPLQASGIAFR	1	13…32
2243.2076		Aa		2243.0627	KPIDRASSHLLISSGYNPK	2	639…657
2485.2880		Aa	2485.2186 or 2485.3971		YDYGGLTLVEMIRSHDSYLPR or LQYVLPLWETNPKFRKPLDK		658…678 or 624…643

a. A list of the most common monoisotopic m/z values from MALDI-TOF analyses of tryptic digests of each HeLp subunit for each tick species (*Amblyomma americanum* (Aa), *A. maculatum* (Am), *Dermacentor andersoni* (Da), *D. variabilis* (Dv), *Ixodes scapularis* (Is), *Rhipicephalus sanguineus* (Rs)) was compiled. *In silico* digests of the region of *R. appendiculatus* or *D. variabilis* HeLp precursor that corresponds to HeLp-A and the region that corresponds to HeLp-B were compared to the list of compiled m/z values for HeLp-A and HeLp-B, respectively, for each species. After residue 429 of HeLp-A, there is a one-residue gap in the alignment of *D. variabilis* HeLp-A with *R. appendiculatus* HeLp-A. For *D. variabilis* HeLp-A after this residue, subtract one from the "start…end" value to obtain residue numbers. The *R. appendiculatus* HeLp-B fragmented sequence was aligned with the *D. variabilis* HeLp-B complete sequence derived from the precursor to obtain residue numbers for *R. appendiculatus* HeLp-B. Differences between the theoretical and measured m/z values were within the mass spectrometer machine mass accuracy for external standards.

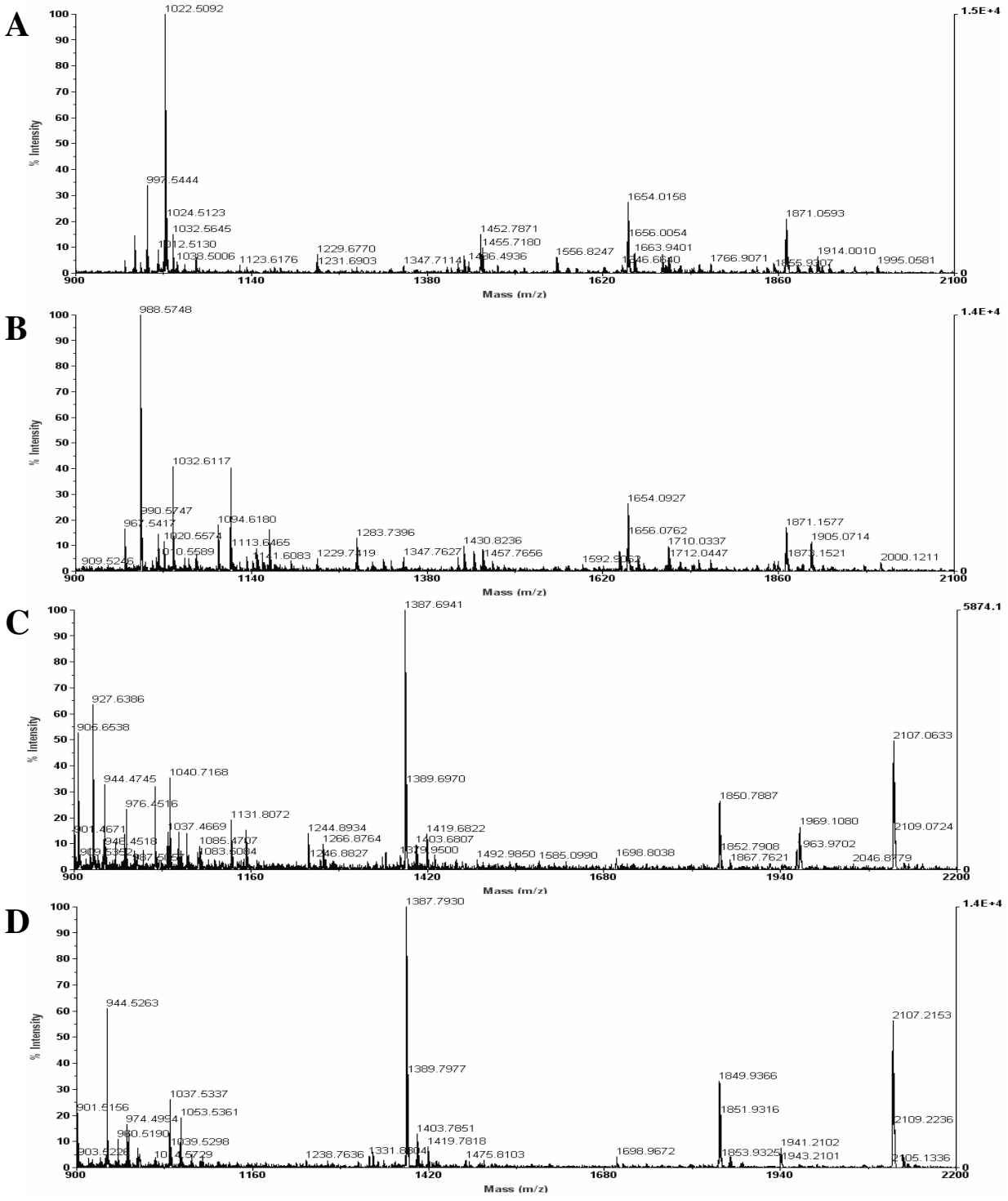


Fig. 2. Typical MALDI-TOF MS traces of peptide mass fingerprints of tryptic digests of HeLp-A from *Dermacentor andersoni* (A) and *Dermacentor variabilis* (B) and HeLp-B from *D. andersoni* (C) and *D. variabilis* (D). For each HeLp subunit, each trace appears similar between both species. Although the *D. andersoni* and *D. variabilis* HeLp subunits share the most monoisotopic m/z values from the HeLp master peak lists, they do not have identical spectra for each subunit.

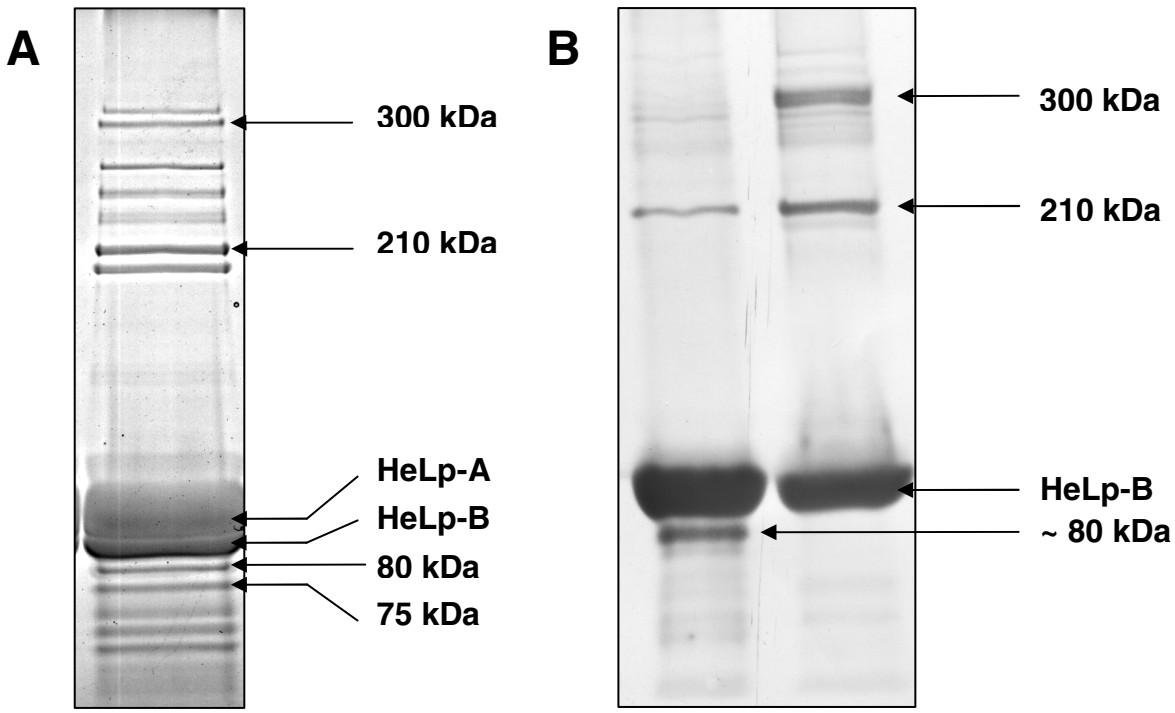


Fig. 3. A. SDS PAGE of *Amblyomma americanum* hemolymph under reducing conditions. Marked bands correspond to the proteins examined for relationship to HeLp. B**. Western analysis of *A. americanum* hemolymph (left) and saliva (right). The primary antibodies (polyclonal) were purified from rabbit serum against synthetic HeLp-B N-terminal peptides (FEVGKEYVYVKYKGTL). In some instances the 300 kDa and/or the ~80 kDa bands appeared with hemolymph samples.**

Table 5: Peak lists containing the most common monoisotopic *m/z* values for *A. americanum* HeLp-A and HeLp-B were generated. A number of peptides separated by SDS PAGE performed under both reducing and non-reducing conditions were digested with porcine trypsin and analyzed by MALDI-TOF mass spectrometry. The resulting fingerprints of all peptides examined showed shared *m/z* values with HeLp-A, HeLp-B, or both.

Peptide Fragment	Reducing Conditions						Non-Reducing Conditions				
m/z (mi)	HeLp-A	HeLp-B	210 kDa	300 kDa	80 kDa	75 kDa	80 kDa	275 kDa	NR-1	NR- 2	NR- 3
974.5279	X		X	X	X	X	X		X	X	X
979.5529	X		X	X	X	X	X		X	X	X
1024.5274	X		X	X		X	X		X		X
1107.5950	X		X	X	X	X	X		X	X	X
1117.6079	X						X		X		X
1134.6283	X		X	X	X	X	X		X	X	X
1282.7254	X		X	X	X	X	X		X	X	X
1374.7293	X		X	X	X		X		X		X
1421.7643	X						X				
1455.7121	X		X	X			X		X	X	X
1475.8524	X								X	X	X
1555.9809	X		X	X	X	X	X		X	X	X
1591.9896	X						X				
1653.9515	X										
1709.9906	X		X	X		X	X				X
1755.9875	X		X			X	X				
1771.9473	X										
1787.9921	X										X
1834.9928	X		X	X	X	X	X			X	X
1855.9843	X		X	X	X	X	X		X	X	X
1869.9468	X										
1890.9648	X						X				
1914.7518	X						X				
1931.0648	X		X	X	X	X	X		X	X	X
1954.1119	X					X	X				X
2022.1326	X		X	X	X	X	X		X	X	X
2038.1687	X										
2054.1376	X										
2570.3920	X						X		X		X
913.5134		X	X	X	X	X		X	X	X	X
945.5111		X									
1008.4853		X	X	X	X			X	X	X	X
1011.5000		X	X	X	X			X	X	X	X
1027.4956		X	X					X	X	X	X
1078.5145		X	X	X	X	X		X	X	X	X
1314.6428		X						X	X		X
1331.8375		X									
1385.7851		X			X			X			X
1423.6547		X				X		X			X
1440.6150		X	X	X	X	X		X	X	X	X
1454.6668		X									
1457.7430		X				X		X	X		X
1471.7650		X			X			X			X
1487.7742		X						X			X
1499.7899		X	X		X			X	X		X
1513.8104		X				X					
1529.8201		X									
1744.8891		X	X	X	X	X		X	X	X	X
1758.9092		X				X					
1865.9409		X			X			X	X	X	X
2186.1622		X	X	X	X	X		X	X	X	X
2243.2076		X			X	X		X	X	X	X
2485.2880		X	X	X	X	X		X			X
2544.3611		X						X			X

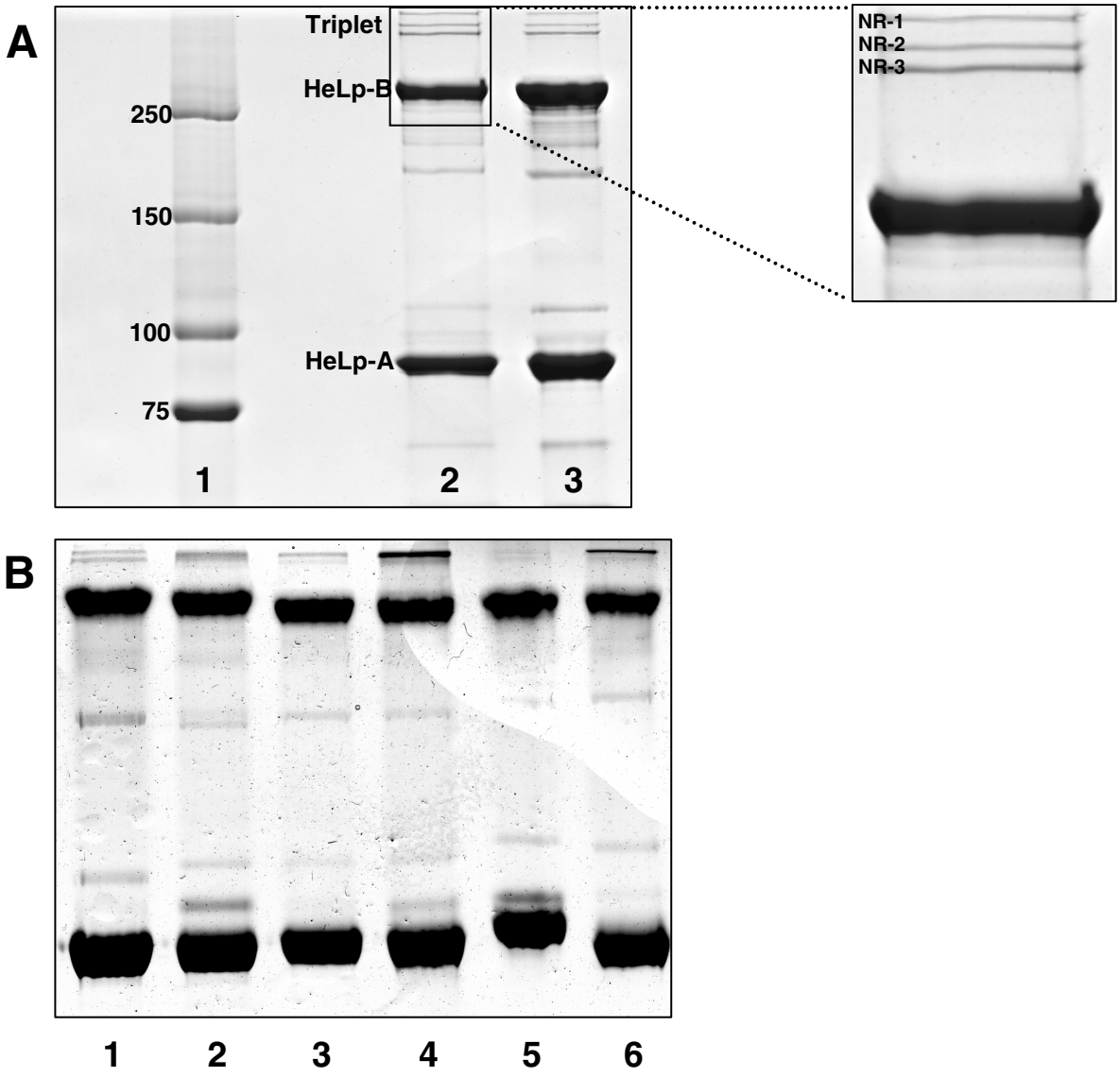


Fig. 4 A. 6.5% SDS PAGE of *A. americanum* saliva (2) and hemolymph (3) proteins under non-reducing conditions. Two major bands migrated to positions corresponding to 275 and 80 kDa. MALDI-TOF analysis revealed these proteins to be solely HeLp-B and HeLp-A, respectively. A high molecular weight triplet reacted with anti-HeLp-B N-terminus. MALDI-TOF MS reveals that all three bands contain peptides of both HeLp-A and HeLp-B. B. In a similar manner, *A. maculatum* (2), *D. andersoni* (3), *D. variabilis* (4), *I. scapularis* (5), and *R. sanguineus* (6) hemolymph proteins were separated by SDS-PAGE under non-reducing conditions, and a banding pattern similar to *A. americanum* (1) results. A number of high molecular weight bands also occurs for these species, although the number of bands in each instance is difficult to distinguish.

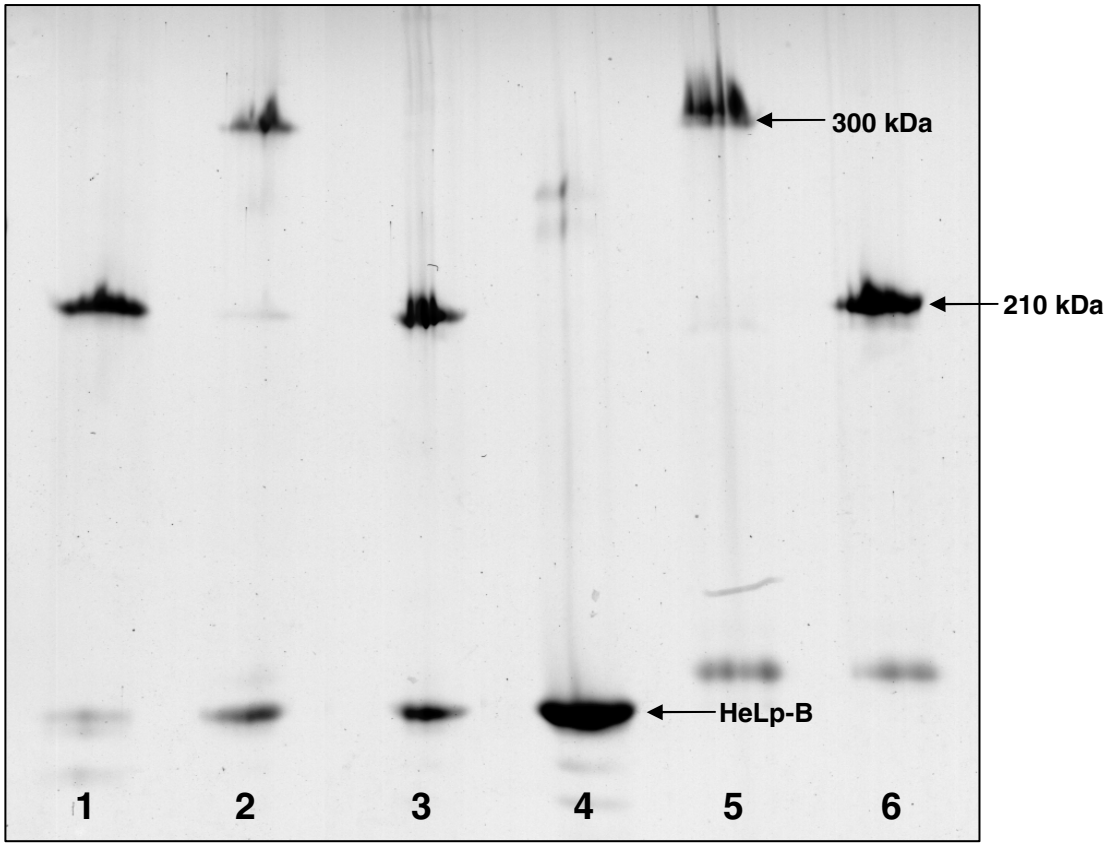


Fig. 5 Reducing SDS-PAGE (6.5%) of bands (*A. americanum*) excised from reduced and non-reduced gels. All proteins were (re)reduced: (1) NR-1, (2) NR-2, (3) NR-3, (4) HeLp-B (non-reduced), (5) 300 kDa HeLp-related protein, (6) 210 kDa HeLp-related protein. The reducing of NR-1, NR-2, and NR-3 reveals that each member of the high molecular weight triplet is composed of the 210 kDa HeLp related protein, the 300 kDa HeLp-related protein and HeLp-B, or the 210 kDa protein and HeLp-B, respectively.

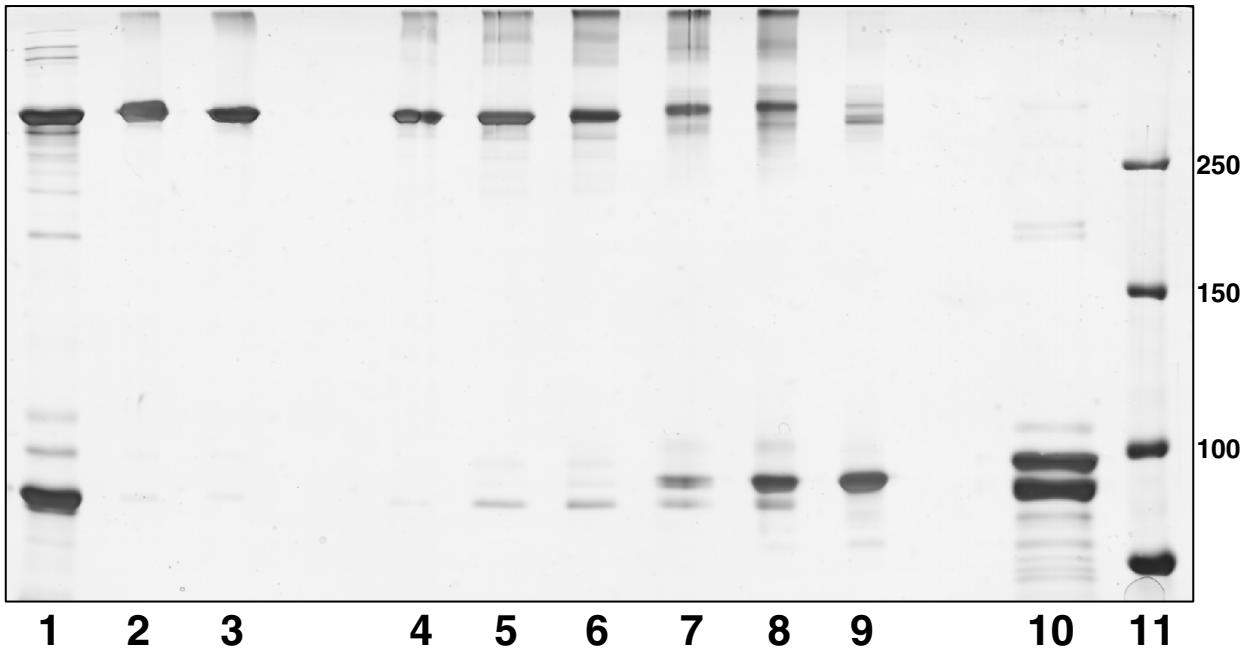


Fig. 6 6.5% SDS-PAGE of a reduction gradient of HeLp-B excised from a non-reduced SDS-polyacrylamide gel. (1) Non-reduced *A. americanum* hemolymph. (2) HeLp-B band, excised from a non-reduced SDS-PAGE analysis of hemolymph. (3) HeLp-B band, excised from a non-reduced SDS-PAGE analysis of hemolymph and subjected to alkylation by iodoacetamide. (4-9) HeLp-B bands, excised from a non-reduced SDS-PAGE *A. americanum* hemolymph gel and subjected to reduction by DTT and alkylation. Gel bands were subjected to reduction in 0.01, 0.05, 0.10, 0.50, 1.0, and 10 mM DTT (left to right) respectively. (10) Reduced *A. americanum* hemolymph. (11) BioRad® Kaleidoscope™ Precision Plus Protein™ standards.

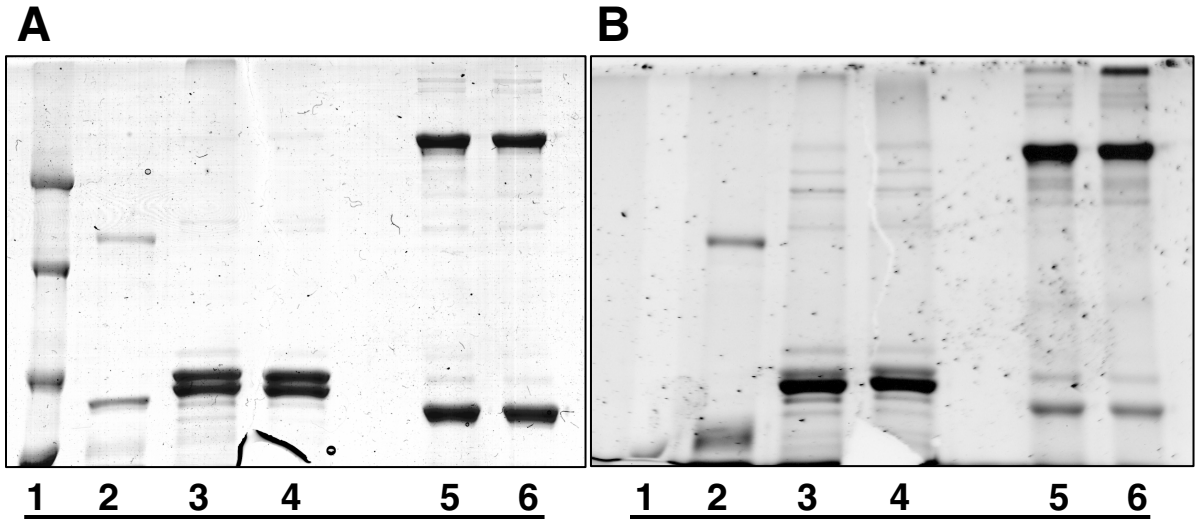


Fig. 7 6.5% SDS-PAGE of *A. americanum* hemolymph and saliva visualized with (A) Coomassie blue R-250 stain and (B) a Pro-Q® Emerald 300 Glycoprotein Gel and Blot Stain Kit. (1) BioRad® Kaleidoscope™ Precision Plus Protein™ standards, (2) CandyCane molecular weight standards (Molecular Probes), (3) reduced *A. americanum* hemolymph, (4) reduced *A. americanum* saliva, (5) non-reduced *A. americanum* hemolymph, (6) non-reduced *A. americanum* saliva.

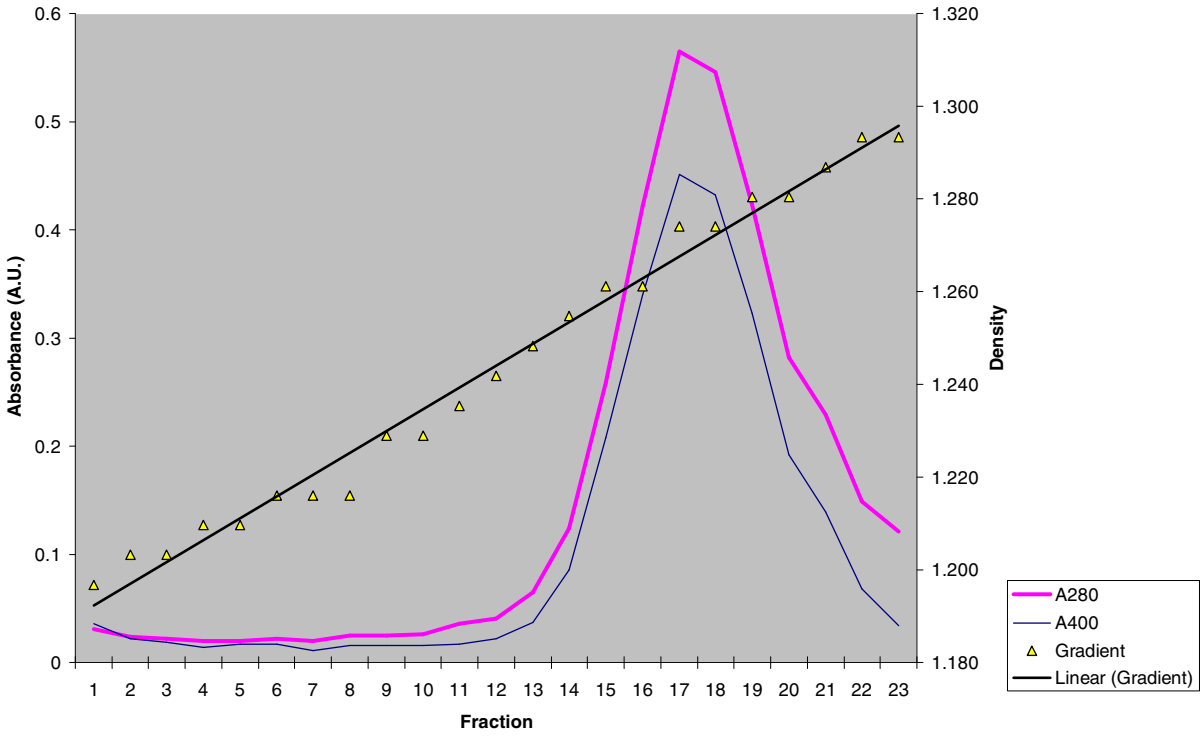


Fig. 8 Fractionation of hemolymph proteins by KBr density gradient ultracentrufugation. HeLp HeLp, the principal hemolymph protein, is isolated using the absorbance at 280 and 400 nm.

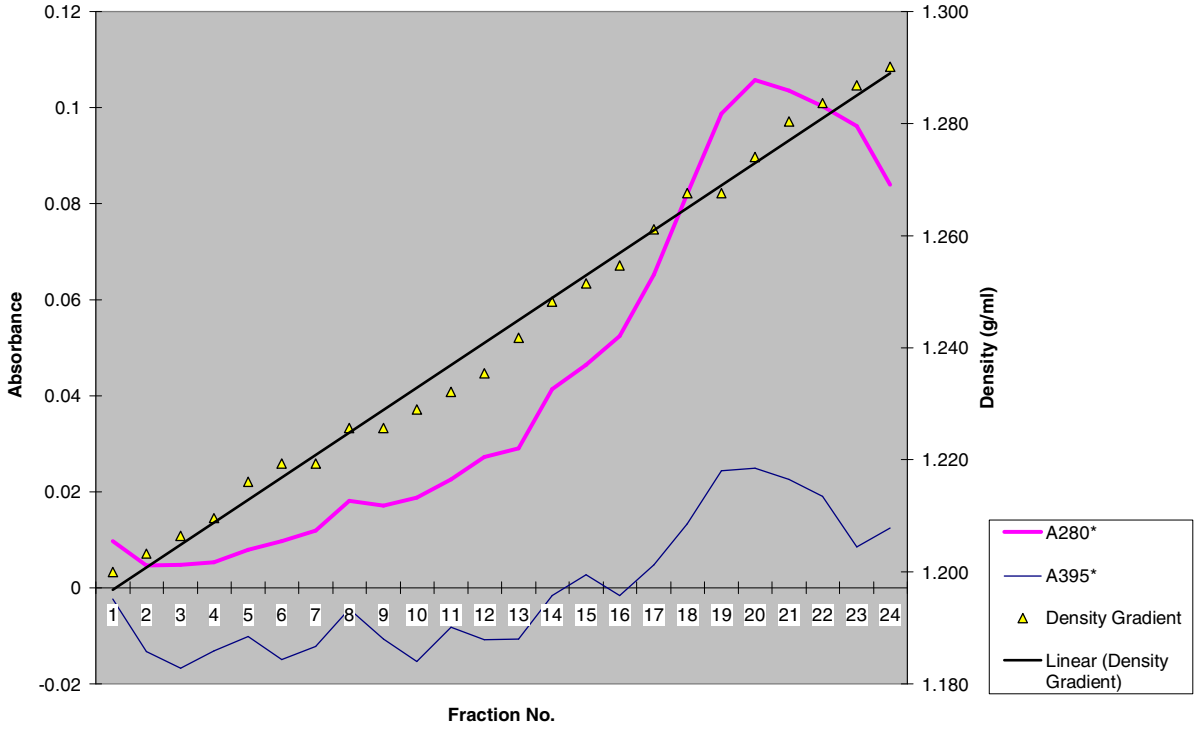


Fig. 9 Fractionation of salivary proteins by KBr density gradient ultracentrufugation. HeLp HeLp, the principal hemolymph protein, is isolated in this instance using the absorbance at 280 and 395 nm.

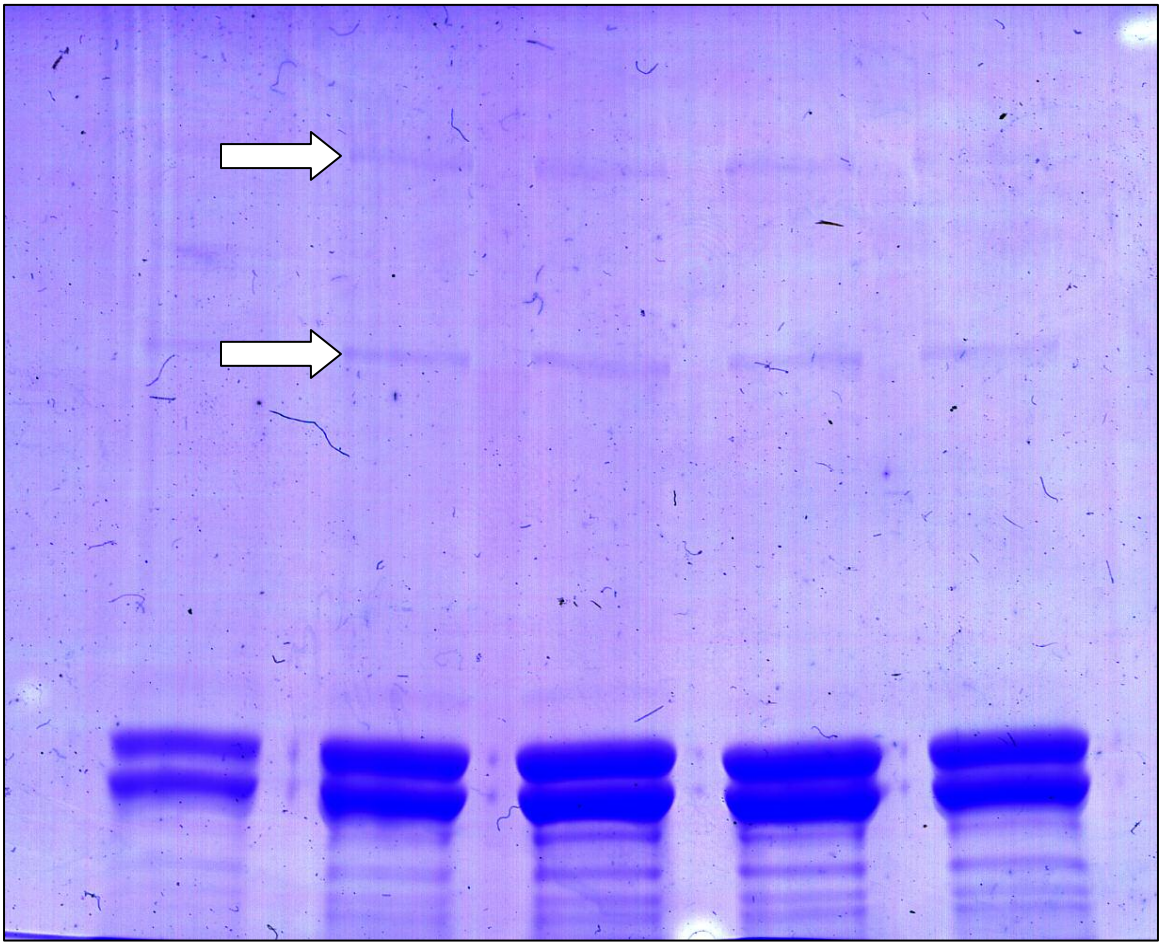
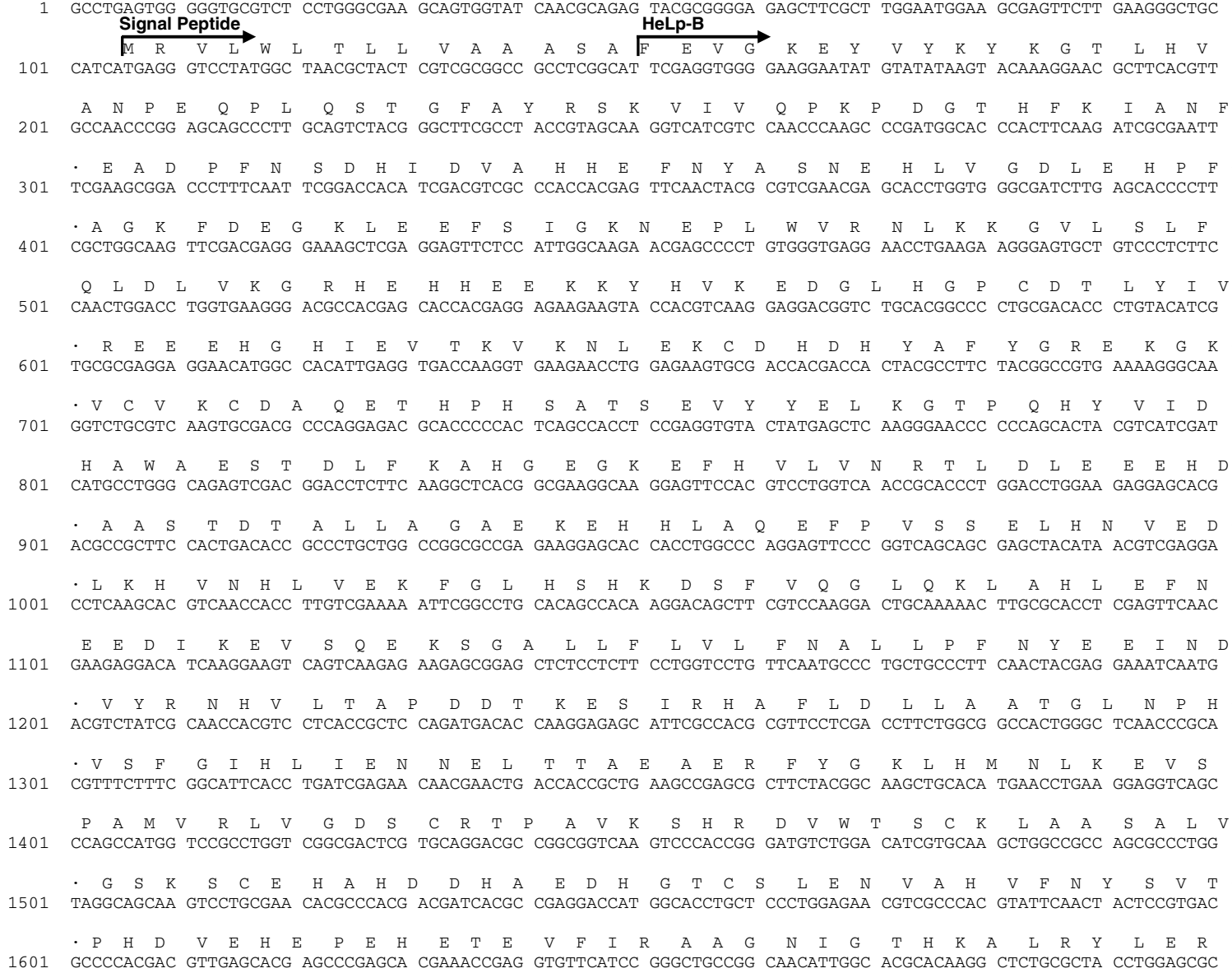


Fig. 10. Fractions 16 through 20 of hemolymph proteins separated by density gradient ultracentrifugation. Arrows are highlighting the faint 300 and 210 kDa bands.


Signal Peptide
HeLp-B
1 GCCTGAGTGG GGGTGCGTCT CCTGGGCGAA GCAGTGGTAT CAACGCAGAG TACGCGGGGA GAGCTTCGCT TGGAATGGAA GCGAGTTCTT GAAGGGCTGC
M R V L W L T L L V A A A S A F E V G K E Y V Y K Y K G T L H V
101 CATCATGAGG GTCCTATGGC TAACGCTACT CGTCGCGGCC GCCTCGGCAT TCGAGGTGGG GAAGGAATAT GTATATAAGT ACAAAGGAAC GCTTCACGTT
A N P E Q P L Q S T G F A Y R S K V I V Q P K P D G T H F K I A N F ·
201 GCCAACCCGG AGCAGCCCTT GCAGTCTACG GGCTTCGCCT ACCGTAGCAA GGTCATCGTC CAACCCAAGC CCGATGGCAC CCACTTCAAG ATCGCGAATT
· E A D P F N S D H I D V A H H E F N Y A S N E H L V G D L E H P F ·
301 TCGAAGCGGA CCCTTTCAAT TCGGACCACA TCGACGTCGC CCACCACGAG TTCAACTACG CGTCGAACGA GCACCTGGTG GGCGATCTTG AGCACCCCTT
· A G K F D E G K L E E F S I G K N E P L W V R N L K K G V L S L F
401 CGCTGGCAAG TTCGACGAGG GAAAGCTCGA GGAGTTCTCC ATTGGCAAGA ACGAGCCCCT GTGGGTGAGG AACCTGAAGA AGGGAGTGCT GTCCCTCTTC
Q L D L V K G R H E H H E E K K Y H V K E D G L H G P C D T L Y I V ·
501 CAACTGGACC TGGTGAAGGG ACGCCACGAG CACCACGAGG AGAAGAAGTA CCACGTCAAG GAGGACGGTC TGCACGGCCC CTGCGACACC CTGTACATCG
· R E E E H G H I E V T K V K N L E K C D H D H Y A F Y G R E K G K ·
601 TGCGCGAGGA GGAACATGGC CACATTGAGG TGACCAAGGT GAAGAACCTG GAGAAGTGCG ACCACGACCA CTACGCCTTC TACGGCCGTG AAAAGGGCAA
· V C V K C D A Q E T H P H S A T S E V Y Y E L K G T P Q H Y V I D
701 GGTCTGCGTC AAGTGCGACG CCCAGGAGAC GCACCCCCAC TCAGCCACCT CCGAGGTGTA CTATGAGCTC AAGGGAACCC CCCAGCACTA CGTCATCGAT
H A W A E S T D L F K A H G E G K E F H V L V N R T L D L E E E H D ·
801 CATGCCTGGG CAGAGTCGAC GGACCTCTTC AAGGCTCACG GCGAAGGCAA GGAGTTCCAC GTCCTGGTCA ACCGCACCCT GGACCTGGAA GAGGAGCACG
· A A S T D T A L L A G A E K E H H L A Q E F P V S S E L H N V E D ·
901 ACGCCGCTTC CACTGACACC GCCCTGCTGG CCGGCGCCGA GAAGGAGCAC CACCTGGCCC AGGAGTTCCC GGTCAGCAGC GAGCTACATA ACGTCGAGGA
· L K H V N H L V E K F G L H S H K D S F V Q G L Q K L A H L E F N
1001 CCTCAAGCAC GTCAACCACC TTGTCGAAAA ATTCGGCCTG CACAGCCACA AGGACAGCTT CGTCCAAGGA CTGCAAAAAC TTGCGCACCT CGAGTTCAAC
E E D I K E V S Q E K S G A L L F L V L F N A L L P F N Y E E I N D ·
1101 GAAGAGGACA TCAAGGAAGT CAGTCAAGAG AAGAGCGGAG CTCTCCTCTT CCTGGTCCTG TTCAATGCCC TGCTGCCCTT CAACTACGAG GAAATCAATG
· V Y R N H V L T A P D D T K E S I R H A F L D L L A A T G L N P H ·
1201 ACGTCTATCG CAACCACGTC CTCACCGCTC CAGATGACAC CAAGGAGAGC ATTCGCCACG CGTTCCTCGA CCTTCTGGCG GCCACTGGGC TCAACCCGCA
· V S F G I H L I E N N E L T T A E A E R F Y G K L H M N L K E V S
1301 CGTTTCTTTC GGCATTCACC TGATCGAGAA CAACGAACTG ACCACCGCTG AAGCCGAGCG CTTCTACGGC AAGCTGCACA TGAACCTGAA GGAGGTCAGC
P A M V R L V G D S C R T P A V K S H R D V W T S C K L A A S A L V ·
1401 CCAGCCATGG TCCGCCTGGT CGGCGACTCG TGCAGGACGC CGGCGGTCAA GTCCCACCGG GATGTCTGGA CATCGTGCAA GCTGGCCGCC AGCGCCCTGG
· G S K S C E H A H D D H A E D H G T C S L E N V A H V F N Y S V T ·
1501 TAGGCAGCAA GTCCTGCGAA CACGCCCACG ACGATCACGC CGAGGACCAT GGCACCTGCT CCCTGGAGAA CGTCGCCCAC GTATTCAACT ACTCCGTGAC
· P H D V E H E P E H E T E V F I R A A G N I G T H K A L R Y L E R
1601 GCCCCACGAC GTTGAGCACG AGCCCGAGCA CGAAACCGAG GTGTTCATCC GGGCTGCCGG CAACATTGGC ACGCACAAGG CTCTGCGCTA CCTGGAGCGC

```
      F  I  S  P  K  W  H  A  N  E  H  E  R  M  A  A  L  W  A  L  K  Q  S  S  R  K  H  P  G  L  A  R  S  I·
1701  TTCATCTCTC CTAAGTGGCA CGCGAATGAG CACGAGAGGA TGGCTGCCCT TTGGGCTCTC AAGCAGTCTT CCCGGAAGCA CCCCGGACTG GCCCGCTCCA

      ·  A  L  P  V  F  H  N  D  S  E  P  S  E  I  R  I  A  A  F  L  V  V  L  V  S  N  P  D  L  Y  I  L  R·
1801  TCGCCCTGCC CGTCTTCCAC AACGACAGCG AGCCCAGCGA GATCCGTATC GCTGCCTTCC TTGTGGTCCT GGTGAGCAAC CCAGATCTGT ACATCCTGCG

      ·  H  I  A  Q  E  V  I  T  D  P  S  D  Q  L  V  A  F  V  T  S  A  F  R  S  M  A  E  S  K  Y  P  C  H
1901  CCACATCGCC CAAGAGGTGA TCACCGACCC CAGCGACCAG TTGGTGGCTT TCGTCACCAG CGCATTCCGC TCGATGGCCG AGTCCAAGTA CCCTTGCCAC

      R  E  I  A  Q  H  L  R  Y  V  L  P  L  W  D  N  I  P  K  L  T  K  P  L  D  K  S  R  S  H  L  T  L  S·
2001  CGTGAAATCG CTCAGCACCT ACGCTACGTT CTGCCTCTGT GGGACAACAT CCCCAAGCTC ACGAAGCCCC TCGACAAGTC CAGGTCTCAC CTGACCCTTT

      ·  S  G  Y  N  P  K  Y  D  F  G  G  A  T  I  M  E  V  I  R  S  H  D  S  Y  F  P  R  N  L  Y  I  N  M·
2101  CCTCCGGATA CAACCCCAAG TACGACTTCG GTGGCGCCAC CATCATGGAG GTGATCCGCT CGCACGACAG CTACTTCCCC CGCAACCTGT ACATCAACAT

      ·  K  D  Y  I  A  G  H  S  T  D  T  L  S  L  S  F  E  S  W  G  M  D  K  L  L  N  K  L  V  G  P  Q  P
2201  GAAGGACTAC ATAGCCGGAC ACTCGACCGA CACTCTCTCC TTATCCTTCG AGAGCTGGGG AATGGACAAG CTGCTGAACA AGCTCGTCGG CCCGCAGCCC
                                  Convertase Site   HeLp-A
      G  S  S  K  S  I  W  D  F  M  G  R  R  R  F  P  R  D  A  S  A  K  E  R  K  E  I  E  D  S  L  H  I  H·
2301  GGATCCAGCA AGAGCATCTG GGACTTCATG GGTCGCCGCA GGTTCCCGCG TGATGCCTCG GCCAAGGAGC GCAAGGAGAT CGAAGACTCC CTTCACATCC

      ·  D  R  E  Y  D  H  A  Y  A  R  L  S  L  S  V  F  G  K  A  I  D  T  W  S  F  D  E  S  I  L  E  K  I·
2401  ATGACCGCGA GTACGACCAC GCCTATGCCC GCCTGAGTCT GTCCGTGTTC GGAAAGGCTA TCGACACCTG GAGCTTCGAT GAATCCATCT TGGAAAAGAT

      ·  K  P  R  D  A  P  E  K  T  A  E  K  L  F  G  Q  E  V  R  K  K  A  F  Y  L  T  Q  D  M  T  Y  L  M
2501  CAAGCCGAGA GACGCCCCAG AGAAGACCGC CGAGAAGCTC TTCGGCCAGG AAGTTCGCAA GAAGGCCTTC TACCTGACCC AGGACATGAC CTACCTCATG

      P  T  E  L  G  V  P  V  F  F  D  F  K  Q  A  E  F  I  Y  A  H  R  H  K  I  D  I  T  H  G  D  S  A  E·
2601  CCCACCGAGC TGGGTGTGCC AGTATTCTTC GACTTCAAGC AGGCCGAGTT CATCTACGCA CACCGCCACA AGATCGACAT CACGCACGGC GACAGCGCGG

      ·  I  N  L  N  I  K  R  H  Y  L  Y  E  V  R  A  Y  Q  M  V  G  F  A  L  T  F  A  Q  S  S  L  G  S  G·
2701  AGATCAACCT GAACATAAAG CGACACTACC TGTACGAAGT TCGCGCCTAC CAGATGGTCG GCTTCGCTCT GACCTTCGCC CAGTCGTCCC TGGGTAGCGG

      ·  Y  D  A  Q  T  L  I  S  W  P  L  D  L  K  A  T  L  A  P  L  E  G  K  L  K  L  H  R  P  L  H  L  P
2801  TTACGACGCC CAGACCCTGA TTTCATGGCC TCTTGATCTG AAGGCGACGC TGGCTCCTCT GGAGGGCAAG CTGAAGCTTC ACCGTCCCCT GCACCTGCCC

      W  N  A  A  N  H  H  F  R  P  F  T  F  Q  M  P  Y  D  L  G  G  D  H  V  N  A  I  T  E  L  S  K  A  Q·
2901  TGGAACGCGG CCAACCACCA CTTCCGCCCG TTCACATTCC AGATGCCCTA CGACCTCGGT GGCGACCACG TCAACGCCAT TACCGAGCTG TCCAAGGCCC

      ·  K  P  L  F  R  A  D  E  L  L  E  F  D  R  H  Y  F  G  D  D  F  G  V  A  M  N  I  K  G  Y  L  V  K·
3001  AGAAGCCGCT TTTCCGTGCC GACGAACTTC TTGAGTTTGA CCGCCACTAC TTCGGTGACG ACTTCGGCGT AGCCATGAAC ATCAAGGGAT ACCTCGTCAA

      ·  K  G  L  H  S  G  L  H  E  F  I  H  E  M  T  A  R  E  R  F  Y  Y  L  L  I  N  P  H  W  H  P  R  D
3101  GAAGGGTCTC CACAGCGGGC TTCACGAGTT CATCCATGAG ATGACTGCGC GGGAGCGTTT CTACTACCTC CTCATCAACC CCCACTGGCA CCCACGTGAC

      V  K  I  Y  F  E  P  A  A  E  D  P  S  T  E  L  D  V  E  I  G  Y  K  F  L  E  H  D  D  E  R  H  S  H·
3201  GTCAAGATTT ACTTCGAGCC GGCCGCTGAA GACCCCTCTA CAGAACTGGA CGTTGAGATC GGCTACAAGT TCCTGGAGCA TGACGACGAA AGGCATAGTC

      ·  F  P  V  H  D  Q  I  G  Q  D  P  E  V  P  S  T  H  V  I  N  L  D  V  S  F  K  G  A  K  E  R  K  V·
3301  ACTTCCCGGT CCACGATCAA ATCGGCCAAG ATCCCGAGGT CCCGTCTACG CACGTGATCA ACCTCGACGT CAGCTTCAAA GGCGCCAAGG AACGCAAGGT
```

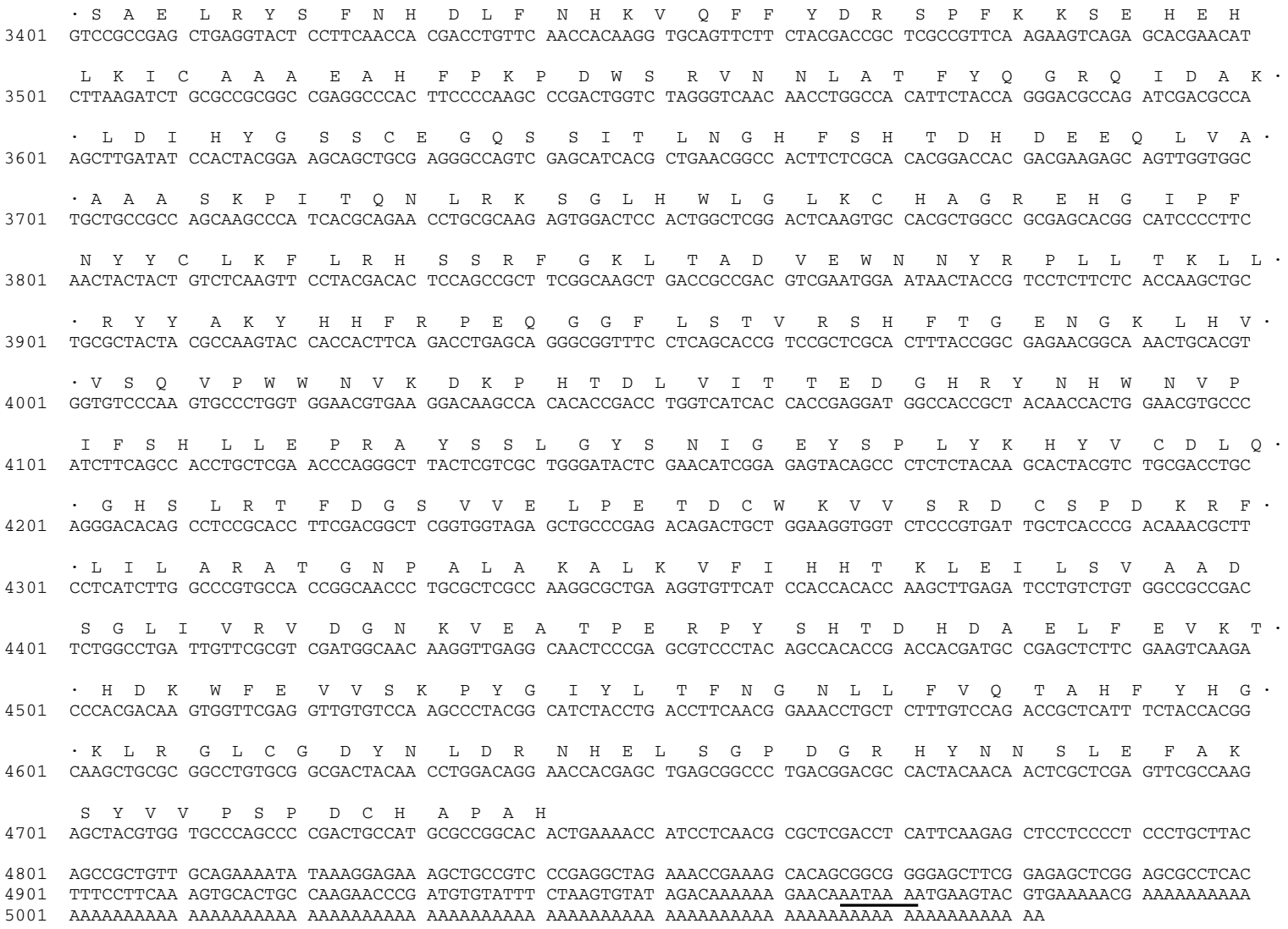


Fig. 11 Contig assembled from AaHeLp cDNA fragments, presumed to represent the whole HeLp precursor transcript, with conceptual translation. Boxed and in bold is the hypothesized dibasic cleavage site, preceding the mature HeLp-A N-terminal sequence. The poly-A signal is underlined.

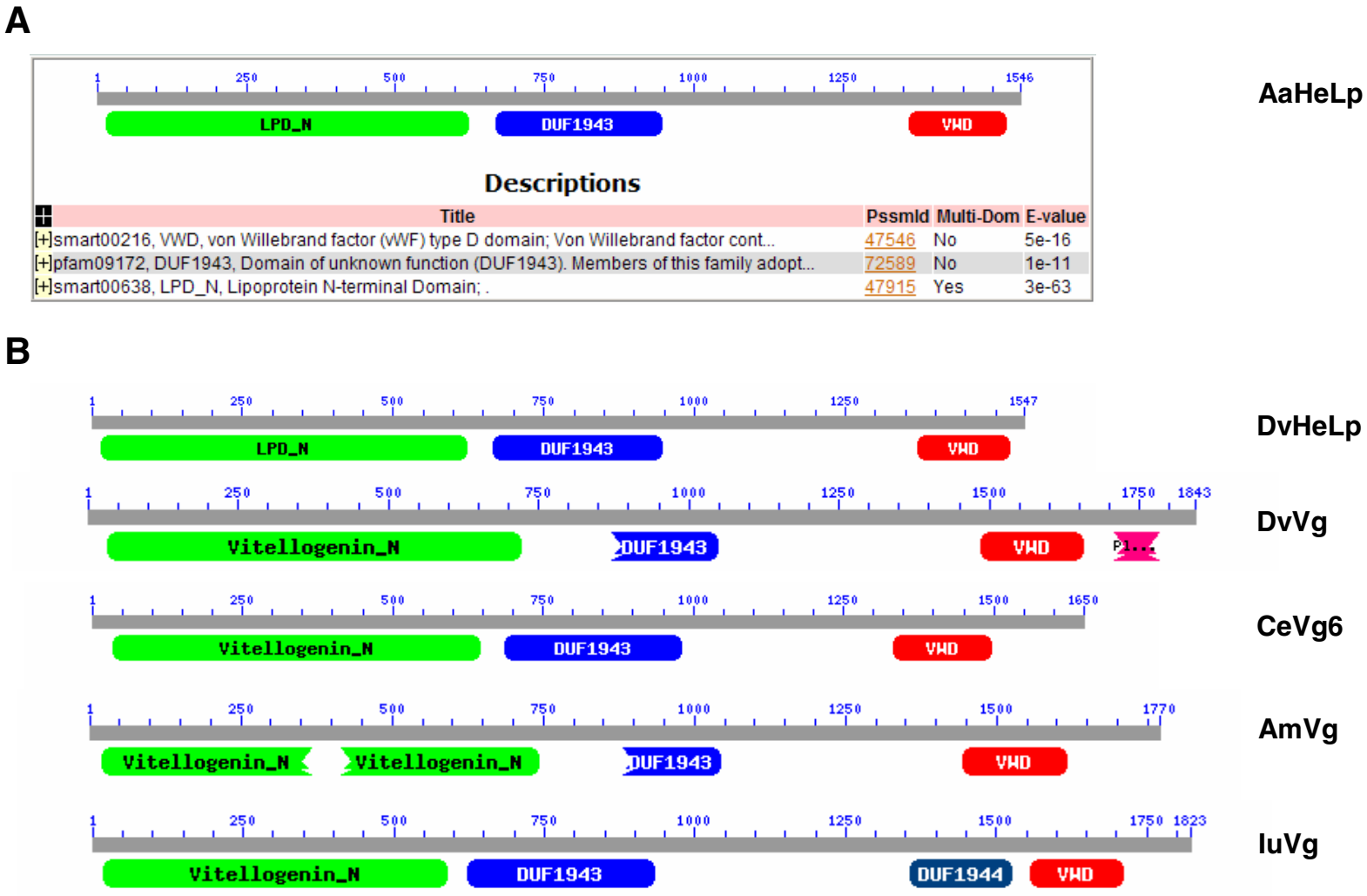


Fig. 12 A comparison of HeLp and its homologues using the Conserved Domain Search at NCBI. (A) AaHeLp precursor domain structure, compared to (B) that of DvHeLp precursor and vitellogenin from a number of species: *D. variabilis*, *Caenorhabditis elegans*, *Apis mellifera*, and *Ichthyomyzon unicuspis*. The N-terminal lipid-binding domain (LPD_N) and the N-terminal vitellogenin domain (Vitellogenin_N) are synonymous domain names. Also pictured are two domains of unknown function (DUF 1943 and DUF 1944) and a von Willebrand factor type D domain (VWD). Figure adapted from output of Conserved Domain Search at http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi.

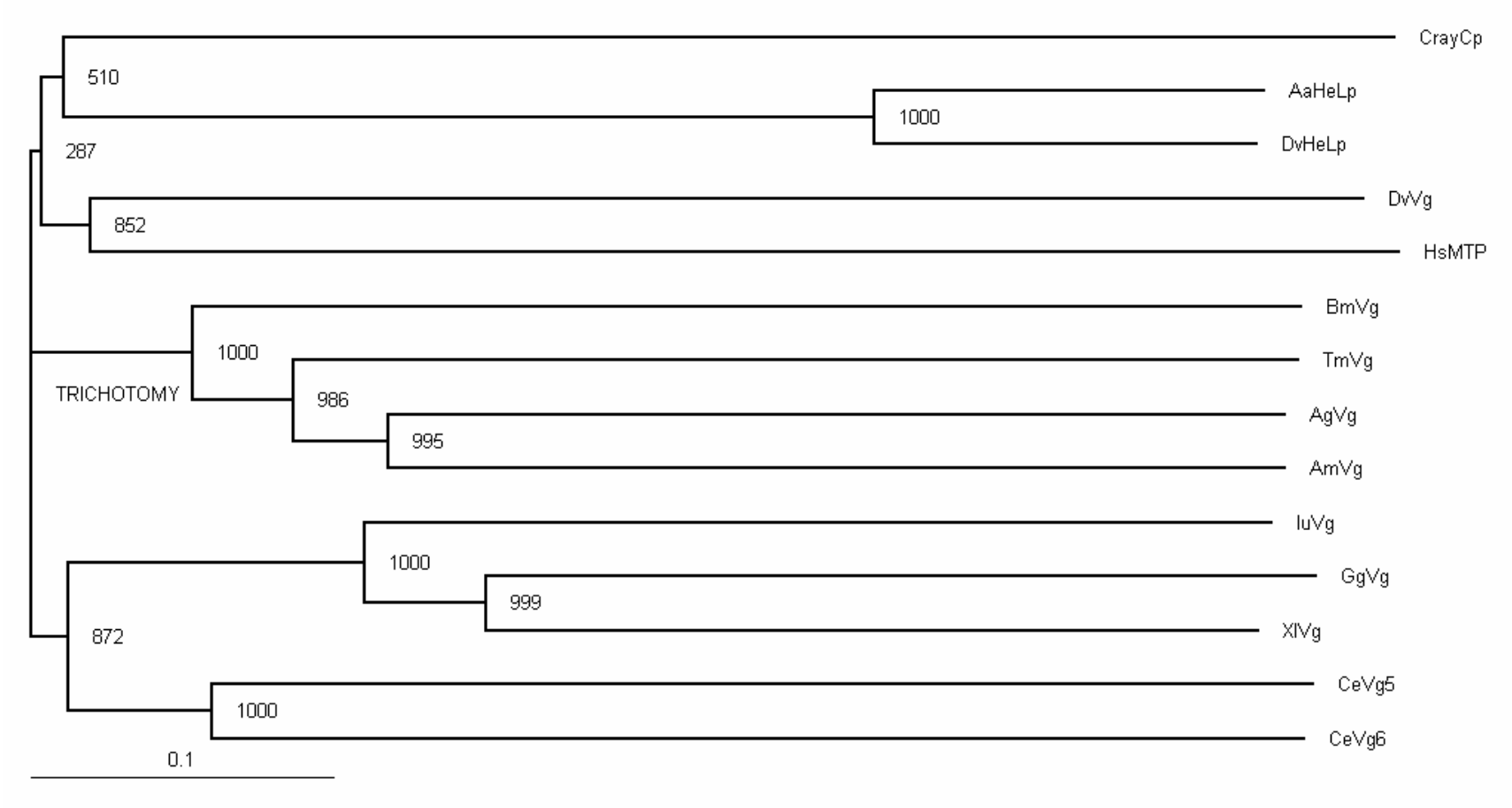


Fig. 13 Unrooted tree constructed using the neighbor-joining method and based on an alignment of the protein sequences of the amino terminus lipid binding domain. The proteins are, in order of appearance, crayfish clotting protein, AaHeLp, DvHeLp, *D.variabilis* vitellogenin, human microsomal triglyceride particle transfer protein, *Bombyx mori* vitellogenin, *Tenebrio molitor* vitellogenin, *Anopheles gambiae* vitellogenin, *Apis mellifera* vitellogenin, *Ichthyomyzon unicuspis* vitellogenin, *Gallus gallus* vitellogenin, *Xenopus laevis* vitellogenin, *Caenorhabditis elegans* vitellogenin-5, and *Caenorhabditis elegans* vitellogenin-6.

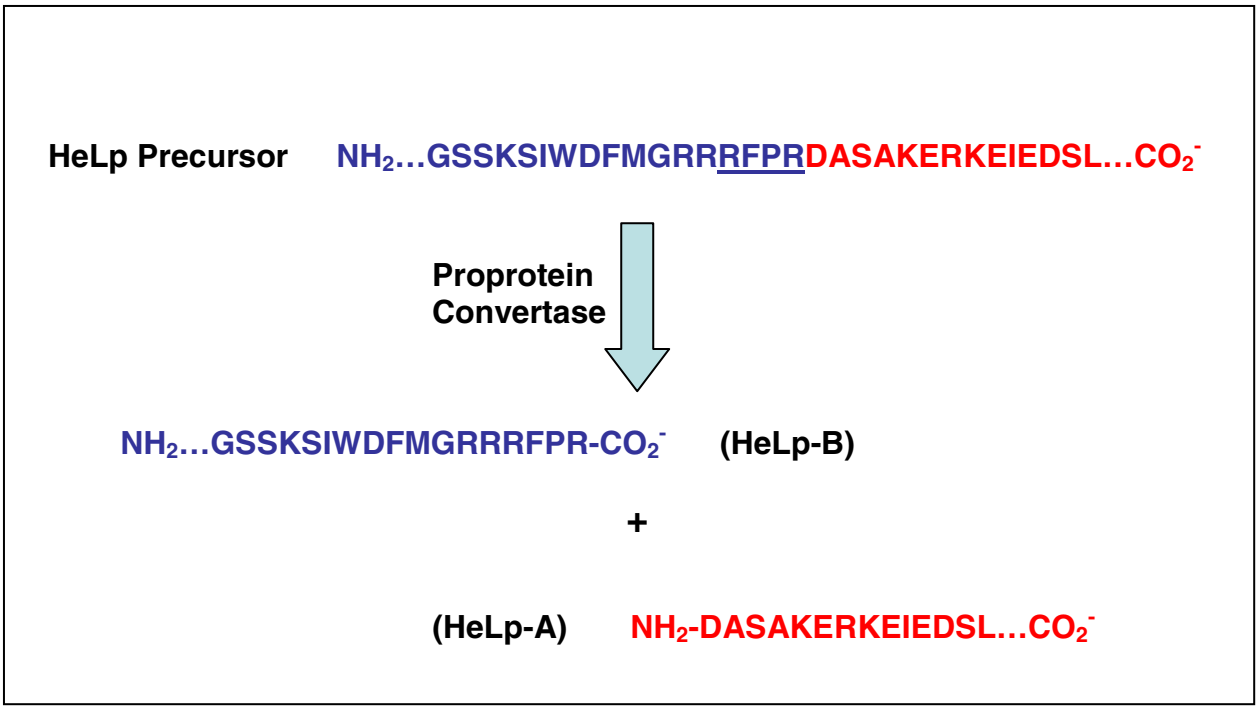


Fig. 14 Proposed processing of the HeLp precursor in the Golgi body.

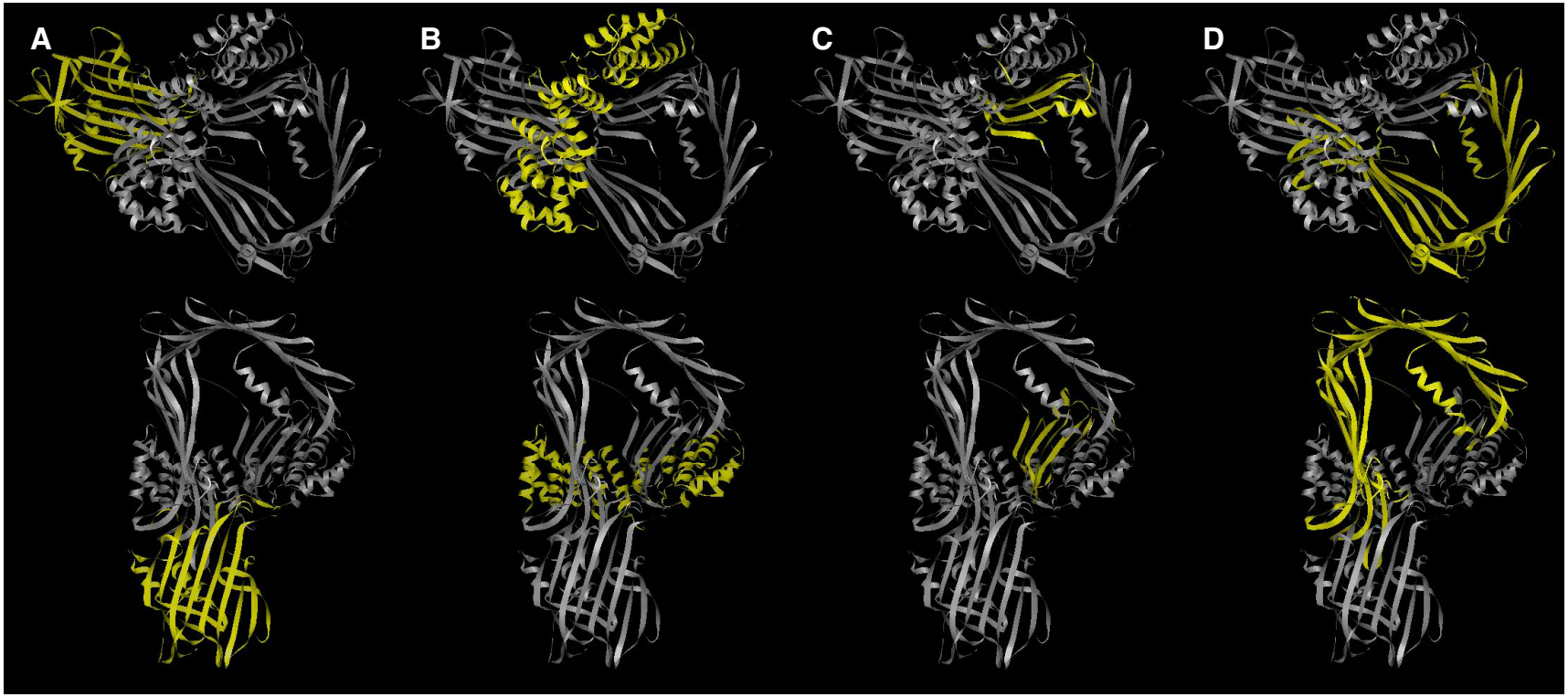


Fig. 15 Ribbon diagram of the (incomplete) lamprey lipovitellin crystal structure, shown with two perspectives. Beginning from the N-terminus, in column A, the N-sheet domain is highlighted; B, the helical domain; C, the C-sheet domain; and D, the A-sheet domain. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0.

```
CLUSTAL X (1.81) multiple sequence alignment


lamprey         MWKLLLVALAFALADAQFQPGKVYRYSYDAFSISGLPEPGVNRAGLSGEMKIEIHGHTHN
AaHeLp          --MRVLWLTLLVAAASAFEVGKEYVYKYKGTLHVANPEQPLQSTGFAYRSKVIVQ-PKPD
                    :*    :. * : *: ** * *.*..    . **  :: :*:: . *: ::  . :

lamprey         QATLKITQVNLKYFLGPWPSDSFYPLTAG-YDHFIQQLEVPVRFDYSAGRIGDIYAPPQV
AaHeLp          GTHFKIANFEADPFNSDHIDVAHHEFNYASNEHLVGDLEHPFAGKFDEGKLEEFSIGKNE
                 : :**::.: . * .   . :.: :. .  :*:: :** *.  .:. *:: ::    :

lamprey         TDTAVNIVRGILNLFQLSLKK------NQQTFELQETGVEGICQTTYVVQEGYRTNEMAV
AaHeLp          PLWVRNLKKGVLSLFQLDLVKGRHEHHEEKKYHVKEDGLHGPCDTLYIVRE-EEHGHIEV
                .  . *: :*:*.****.* *      :::.:.::* *:.* *:* *:*:*  . ..: *

lamprey         VKTKDLNNCDHKVYKTMGTAYAERCPTCQKMNKN-LRSTAVYNYAIFDEPSGYIIKSAHS
AaHeLp          TKVKNLEKCDHDHYAFYGREKGKVCVKCDAQETHPHSATSEVYYELKGTPQHYVIDHAWA
                .*.*:*::***. *   *   .: * .*:  :.:   :*:   * : . *. *:*. * :

lamprey         EEIQQLSVFDIK-EGNVVIESRQKLILEGIQSAPAASQAASLQNRGGLMYKFPSS-AITK
AaHeLp          ESTDLFKAHGEGKEFHVLVNRTLDLEEEHDAASTDTALLAGAEKEHHLAQEFPVSSELHN
                *. : :....   * :*:::   .*  *   ::. ::  *. ::.  *  :** *  : :

lamprey         MSSLFVTKGKNLESEIHTVLKHLVE--------------NNQLSVHEDAPAKFLRLTAFL
AaHeLp          VEDLKHVNHLVEKFGLHSHKDSFVQGLQKLAHLEFNEEDIKEVSQEKSGALLFLVLFNAL
                :..*  .:    :  :*:  . :*:               :::* .:...  ** *   *

lamprey         RNVDAGVLQSIWHKLHQ------QKDYRRWILDAVPAMATSEALLFLKRTLASEQLTSAE
AaHeLp          LPFNYEEINDVYRNHVLTAPDDTKESIRHAFLDLLAATGLNPHVSFGIHLIENNELTTAE
                  .:   ::.::::         ::. *: :** :.* . .  : *  : : .::**:**

lamprey         ATQIVYSTLSNQQATRESLS-YARELLHTSFIRNRPILRKTAVLGYGSLVFR--------
AaHeLp          AERFYGKLHMNLKEVSPAMVRLVGDSCRTPAVKSHRDVWTSCKLAASALVGSKSCEHAHD
                * ::  .   * : .  ::   . :  :*. ::.:  : .:. *. .:**

lamprey         -YCANTVSCPDELLQPLHDLLSQSSDRADE--EEIVLALKALGNAGQPNSIKKIQRFLPG
AaHeLp          DHAEDHGTCSLENVAHVFNYSVTPHDVEHEPEHETEVFIRAAGNIGTHKALRYLERFISP
                 :. :  :*. * :  :.:    . *  .*  .*  : ::* ** *  :::: ::**:.

lamprey         QGKSLDEYSTRVQAEAIMALRNIAKRDPRKVQEIVLPIFLNVAIKSELRIRSCIVFFESK
AaHeLp          KWHAN----EHERMAALWALKQSSRKHPGLARSIALPVFHNDSEPSEIRIAAFLVVLVSN
                : ::      : :  *: **:: :::.*  .:.*.**:* * :  **:** : :*.: *:

lamprey         PSVALVSMVAVRLRREPNLQVASFVYSQMRSLSRSSNPEFRDVAAACSVAIKMLG--SKL
AaHeLp          PDLYILRHIAQEVITDPSDQLVAFVTSAFRSMAESKYPCHREIAQHLRYVLPLWDNIPKL
                *.: ::  :* .:  :*. *:.:** * :**::.*. * .*::*     .: : .  .**

lamprey         DR-LGCRYSKAVHVDTFNARTMAGVSADYFRINSPSGPLPRAVAAKIRGQGMGYASDIVE
AaHeLp          TKPLDKSRSHLTLSSGYNPKYDFGGATIMEVIRSHDSYFPRNLYINMKDYIAGHSTDTLS
                 : *.   *: .  . :*.:   * ::    *.* .. :** :  :::.   *:::* :.

lamprey         FGLRAEGLQELLYRGSQEQDAYGTALDRQTLLRSGQARSHVSSIHDTLRKLSDWKSVPEE
AaHeLp          LSFESWGMDKLLNKLVGPQPGS-SKSIWDFMGRRRFPRDASAKERKEIEDSLHIHDREYD
                :.:.: *:::** :    * .  :    : : *   .*.  :. :. :..  . :.   :

lamprey         RPLASGYVKVHGQEVVFAELDKKMMQRISQLWHSARSHHAAAQEQIRAVVSKLEQGMDVL
AaHeLp          HAYARLSLSVFGKAIDTWSFDESILEKIKPRDAPEKTAEKLFGQEVRKKAFYLTQ-----
```

```
                :. *   :.*.*: :   .:*:.::::*.    . :: .    :::*  .  * *

lamprey         LTKGYVVSEVRYMQPVCIGIPMDLNLLVSGVTTNRANLHASFSQSLPADMKLADLLATNI
AaHeLp          --------DMTYLMPTELGVPVFFDFKQ-----------AEFIYAHRHKIDITHGDSAEI
                        :: *: *. :*:*: :::            *.*  :   .:.::.  :::*

lamprey         ELRVAATTSMSQHAVAIMGLTTDLAKAGMQTHYKTSAGLGVNGKIEMNARESNFKASLKP
AaHeLp          NLNIKRHYLYEVRAYQMVGFALTFAQSSLGSGYDAQT---------LISWPLDLKATLAP
                :*.:      . :*  ::*::  :*::.: : *.:.:         : :   ::**:* *

lamprey         FQQKTVVVLSTMESIVFVRDPSGSRILPVLPPKMTLDKGLISQQQQQPHHQQQPHQHGQD
AaHeLp          LEGKLKLHR-----------------PLHLPWNAANHHFRPFTFQMPYDLGGDHVNAIT
                :: *  :                   *:  *  : :: : .   * *:.    * :.

lamprey         QARAAYQRPWASHEFSPAEQKQIHDIMTARPVMRRKQHCSKSAALSSKVCFSARLRNAAF
AaHeLp          ELSKAQKPLFRADELLEFDRHYFGDDFGVAMNIKG--YLVKKGLHSGLHEFIHEMTARER
                :   * :  : :.*:   ::: : * : .   ::   :  *..  *.   *  .:

lamprey         IRNALLYKITGDYVSKVYVQPTSSKAQIQKVELELQAGPQAAEKVIRMVELVAKASKKSK
AaHeLp          FYYLLINPHWHPRDVKIYFEPAAEDP---STELDVEIG----------------------
                :   *:         *:*.:*::...   ..**::: *

lamprey         KNSTITEEGVGETIISQLKKILSSDKDKDAKKPPGSSSSSSSSSSSSSSSSSDKSGKKT
AaHeLp          ----------------YKFLEHDDERHSHFPVHDQIGQDPEVPST-------------
                              *:*. *.::.:: *  .. ..... .*:

lamprey         PRQGSTVNLAAKRASKKQRGKDSSSSSSSSSSSDSSKSPHKHGGAKRQHAGHGAPHLGP
AaHeLp          ----HVINLDVSFKGAKERKVSAELRYSFNHDLFN------------------------
                     .:** ..  . *:*  .:.   * . .  :

lamprey         QSHSSSSSSSSSSSSSASKSFSTVKPPMTRKPRPARSSSSSSSDSSSSSSSSSSSSSS
AaHeLp          --HKVQFFYDRSPFKKSEHEHLKICAAAEAHFPKPDWSRVNNLATFYQGRQIDAKLDIHY
                  *. .   . *. ..*  : :.   .. :: *:*  *  .. ::  .. . .:. .

lamprey         SSSSSSSSSESKSLEWLAVKDVNQSAFYNFKYVPQRKPQTSRRHTPASSSSSSSSSSSS
AaHeLp          GSSCEGQSS----------------------ITLNGHFSHTDHDEEQLVAAAASKPITQ
                .**....**                       :. .   :   *   .  ::::*.. :.

lamprey         SSSSSSDSDMTVSAESFEKHSKPKVVIVLRAVRADGKQQGLQTTLYYGLTSNGLPKAKIV
AaHeLp          NLRKSGLHWLGLKCHAGREHGIPFNYYCLKFLRHSSRFG--------------------
                .  .*.   : :...: .:*. *     *: :* ..:

lamprey         AVELSDLSVWKLCAKFRLSAHMKAKAAIGWGKNCQQYRAMLEASTGNLQSHPAARVDIKW
AaHeLp          ----------------------KLTADVEWNNYRPLLTKLLRYYAKYHHFRPEQ------
                                     * .* : *.:       :*.  :   : :*

lamprey         GRLPSSLQRAKNALLENKAPVIASKLEMEIMPKKNQKHQVSVILAAMTPRRMNIIVKLPK
AaHeLp          ----GGFLSTVRSHFTGENGKLHVVSQVPWWNVKDKPHTDLVITTEDGHRYN--------
                    ..:  : .: : .:   :    ::     *:: *   ** :    *

lamprey         VTYFQQGILLPFTFPSPRFWDRPEGSQSDSLPAQIASAFSGIVQDPVASACELNEQSLTT
AaHeLp          ------------HWNVPIFSHLLEPRAYSSLGYSNIGEYSPLYKH---YVCDLQGHSLRT
                             :  * * .  *    .**  .  . :* : :.    .*:*: :** *

lamprey         FNGAFFNYDMPESCYHVLAQECSSRPPFIVLIKLDSERRIS--LELQLDDKKVKIVSRND
AaHeLp          FDGSVVELPETD-CWKVVSRDCSPDKRFLILARATGNPALAKALKVFIHHTKLEILSVAA
                *:*:..:   .: *::*::::**.   *::* :  .:  ::  *:: :...*::*:*
```

```
lamprey          -----IRVDG---EKVPLRRLSQKNQ----YGFLVLDAGVHLLLKYKDLRVSFNSSSVQV
AaHeLp           DSGLIVRVDGNKVEATPERPYSHTDHDAELFEVKTHDKWFEVVSKPYGIYLTFNGNLLFV
                      :****   * .* *  *:.::    : . . *  ..:: *  .: ::**.. : *

lamprey          WVPSSLKGQTCGLCGRNDDELVTEMRMPNLEVAKDFTSFAHSWIAPDETCGGACALSRQT
AaHeLp           QTAHFYHGKLRGLCGDYNLDRNHELSGPDGRHYNNSLEFAKSYVVPSPDCHAPAH-----
                  ..   :*:  ****  : :   *:  *: .  ::  .**:*::.*.  * ...

lamprey          VHKESTSVISGSRENCYSTEPIMRCPATCSASRSVPVSVAMHCLPAESEAISLAMSEGRP
AaHeLp           ------------------------------------------------------------


lamprey          FSLSGKSEDLVTEMEAHVSCVA
AaHeLp           ----------------------
```

Fig. 16 ClustalW pairwise alignment of lamprey vitellogenin precursor and AaHeLp precursor using a BLOSUM30 scoring matrix.

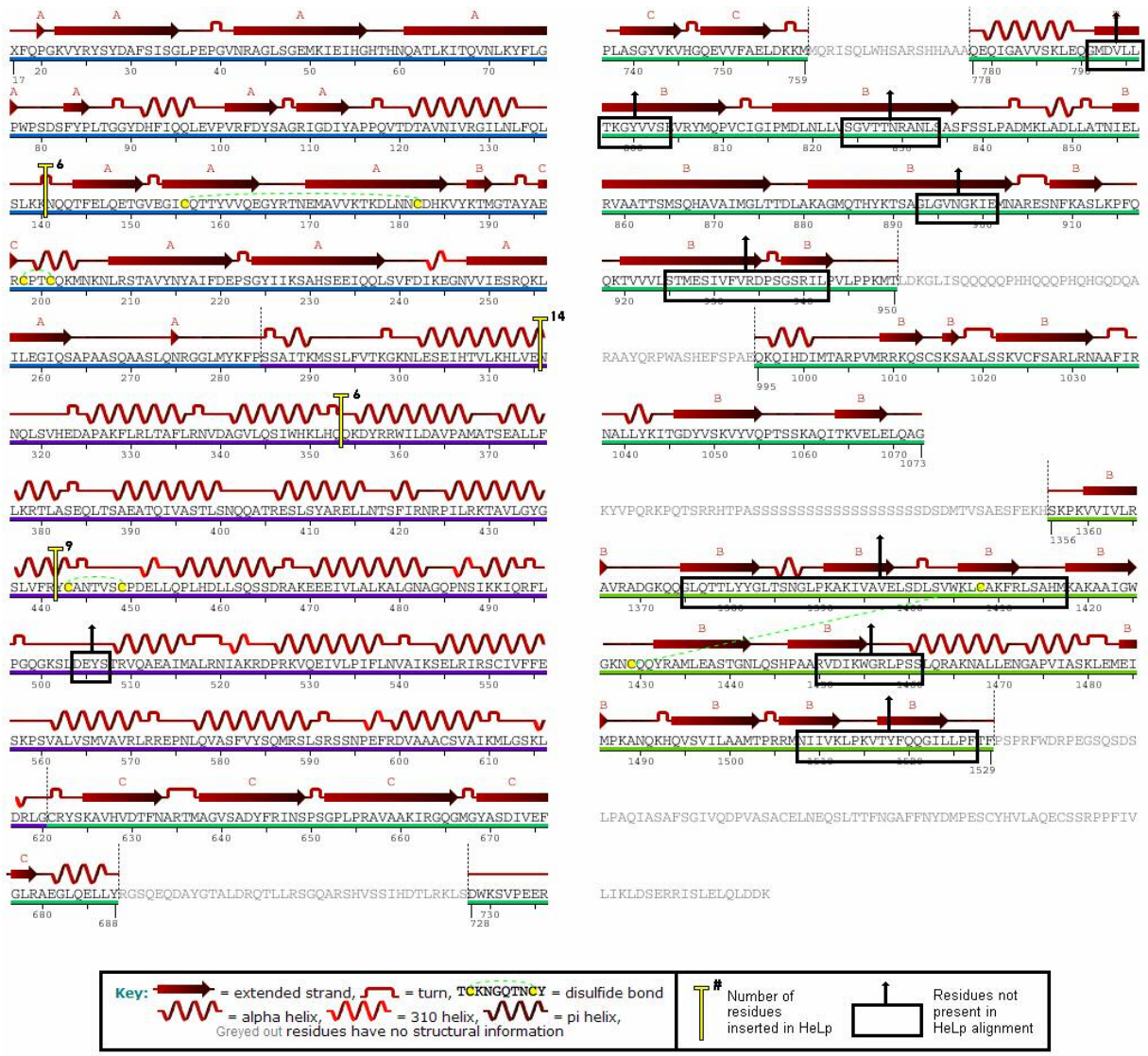


Fig. 17 Actual secondary structure of lamprey lipovitellin (1LSH), modified from the Protein Data Bank at http://www.rcsb.org/pdb/. Depicted are regions that, according to an alignment of lipovitellin and AaHeLp, differ between the two proteins.

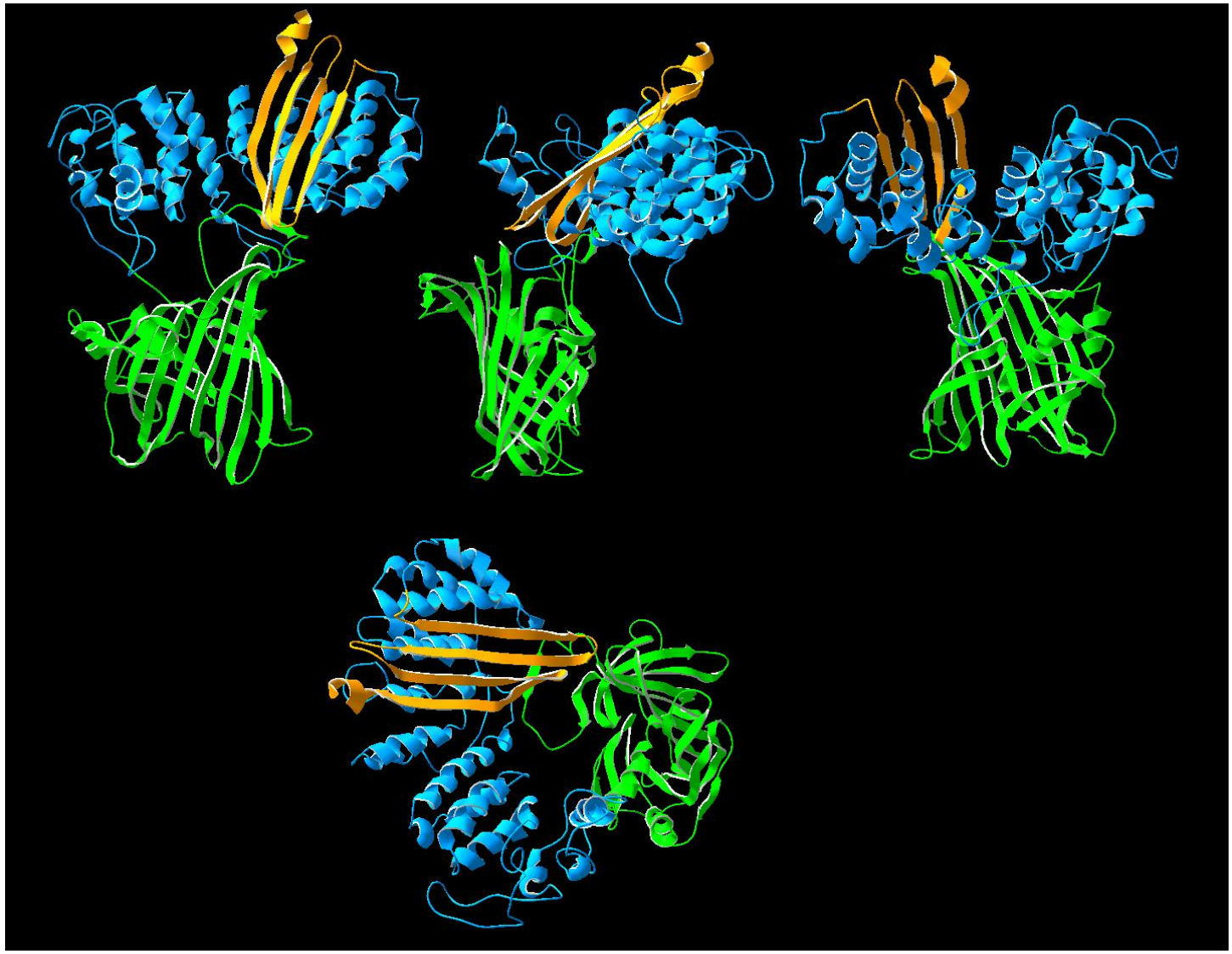


Fig. 18 Ribbon diagram of HeLp-B model made using the CPH 2.0 homology modeler, viewed from different perspectives. Depicted are the green N-sheet, blue helical domain, and orange C-sheet. Figure created with DeepView.

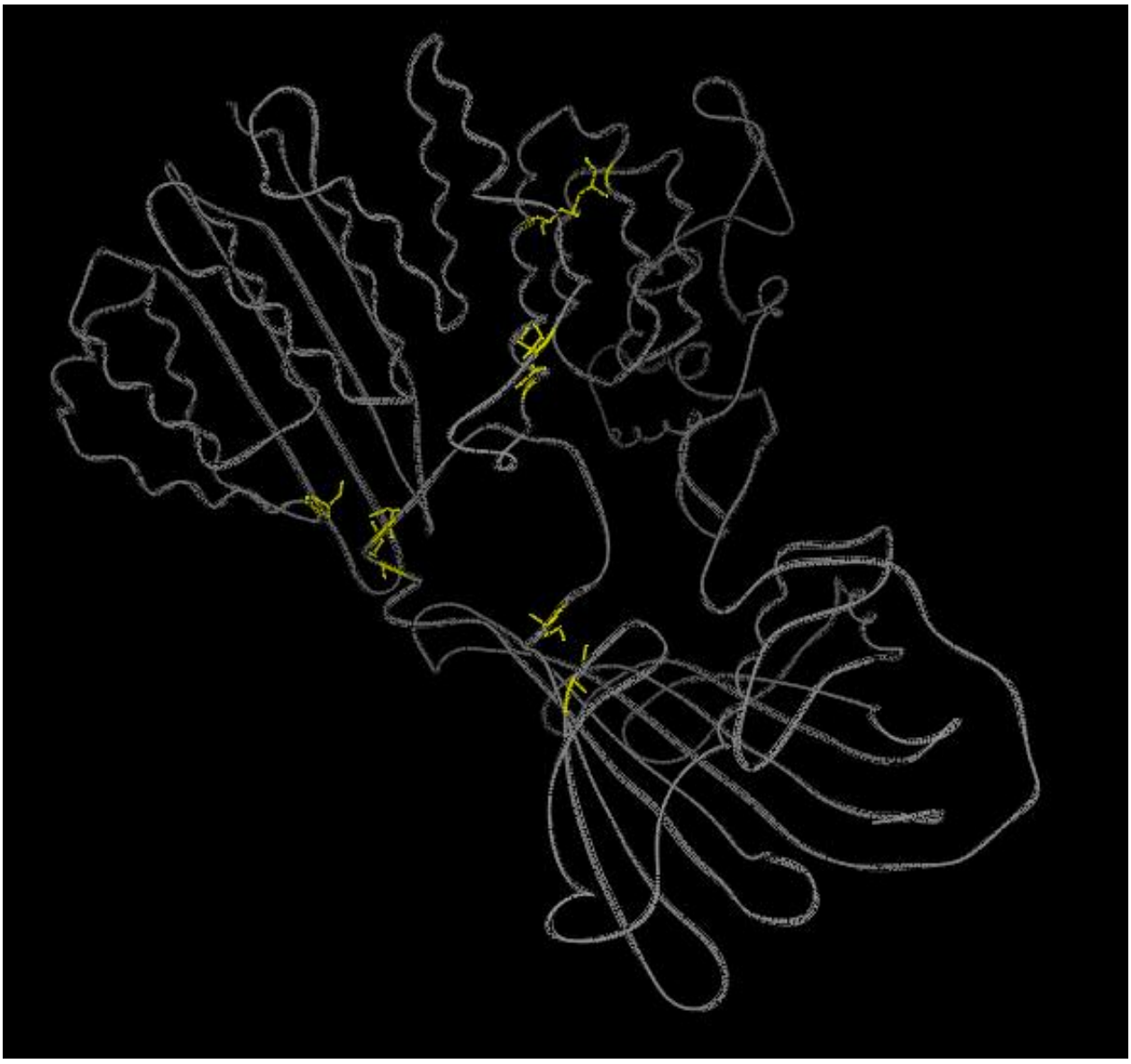


Fig. 19 Relative positions of cysteines on the modeled HeLp-B backbone. Cysteines are highlighted in yellow. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0.

Fig. 20 Positions of C-428 and C-443 in the helical domain of the modeled HeLp-B backbone. Cysteines of interest are highlighted in yellow. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0.

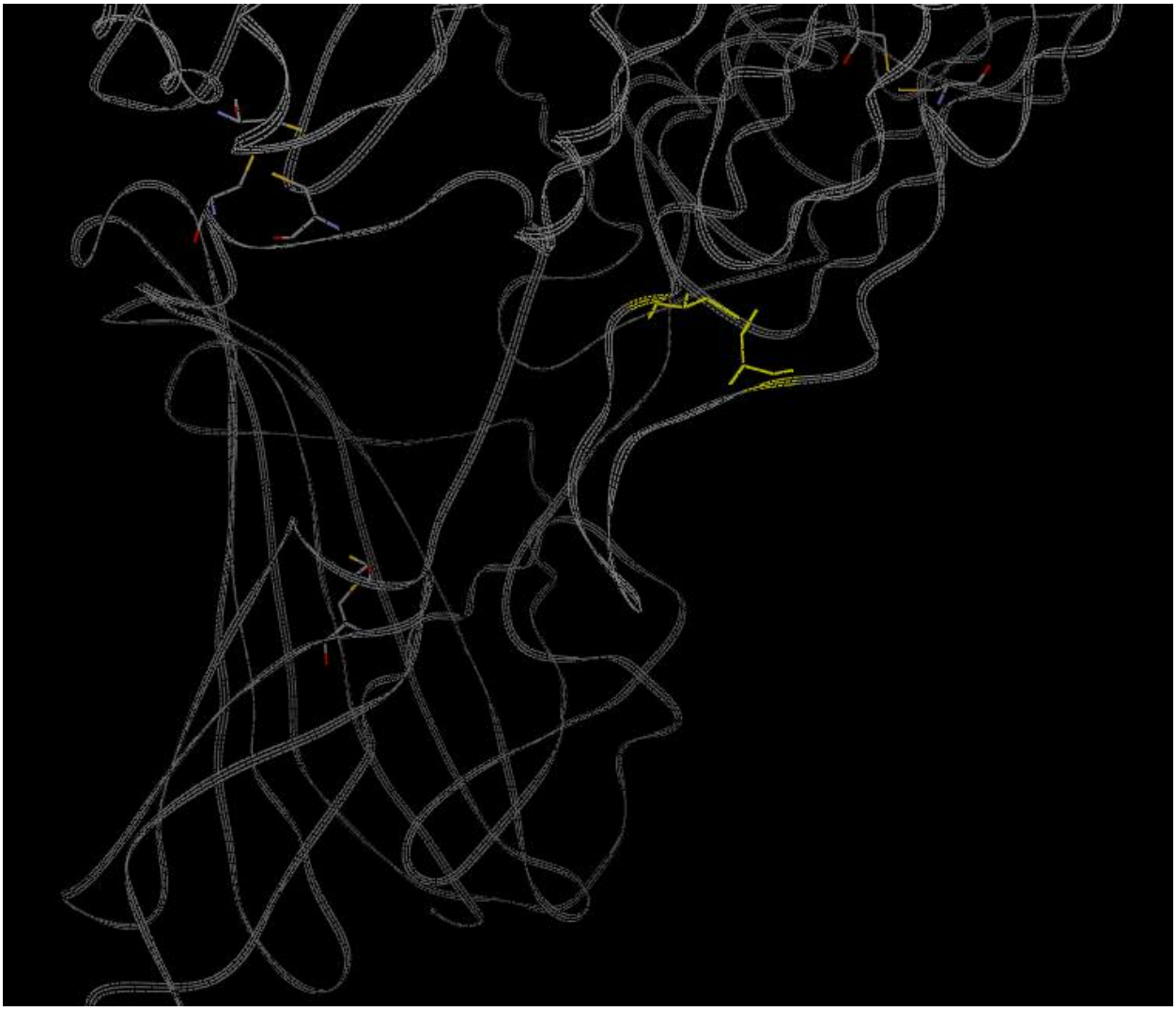


Fig. 21 Positions of C-456 and C-470 in the helical domain of the modeled HeLp-B backbone. Cysteines of interest are highlighted in yellow. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0.

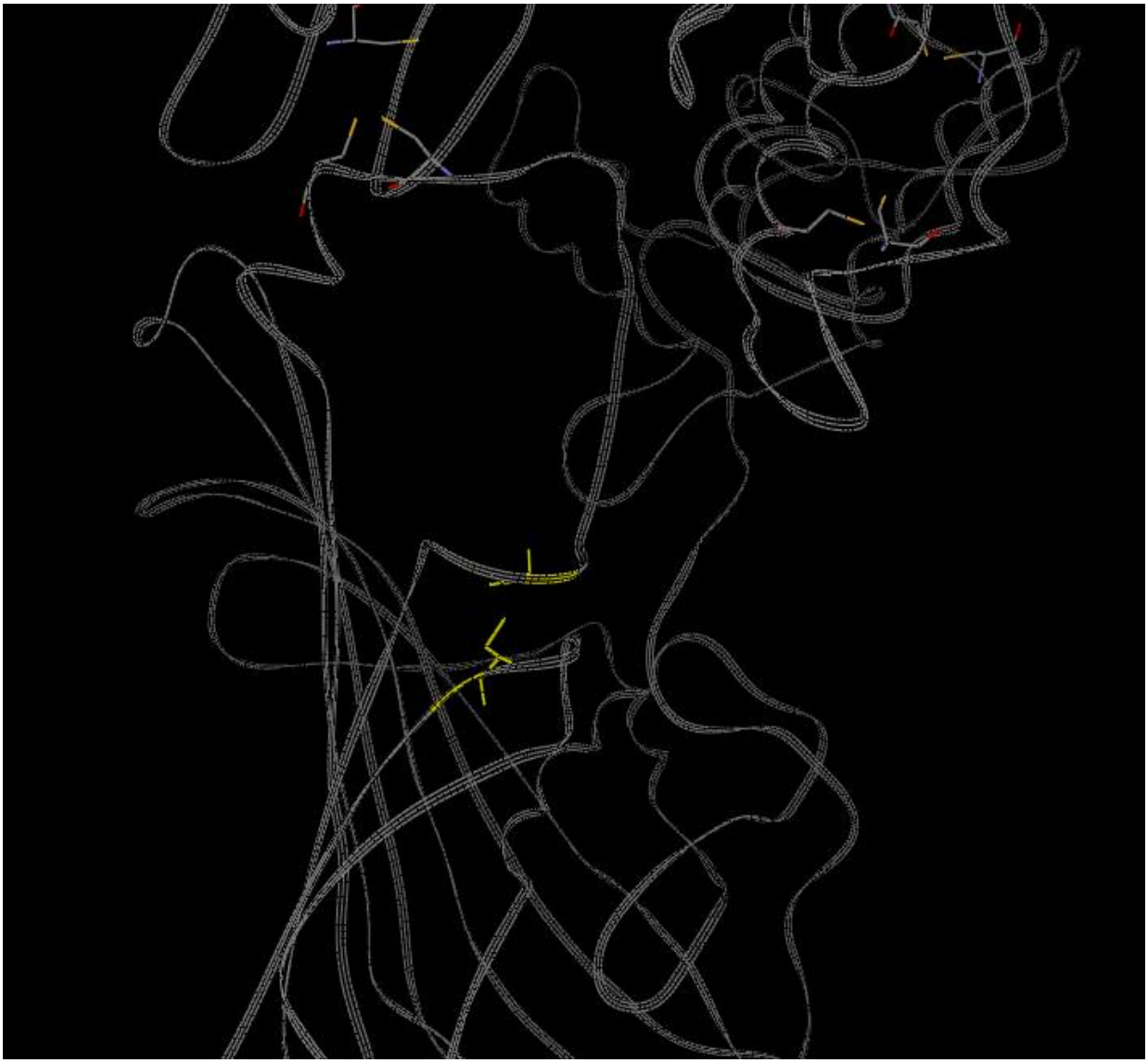


Fig. 22 Positions of C-145 and C-170 in the N-sheet of the modeled HeLp-B backbone. Cysteines of interest are highlighted in yellow. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0.

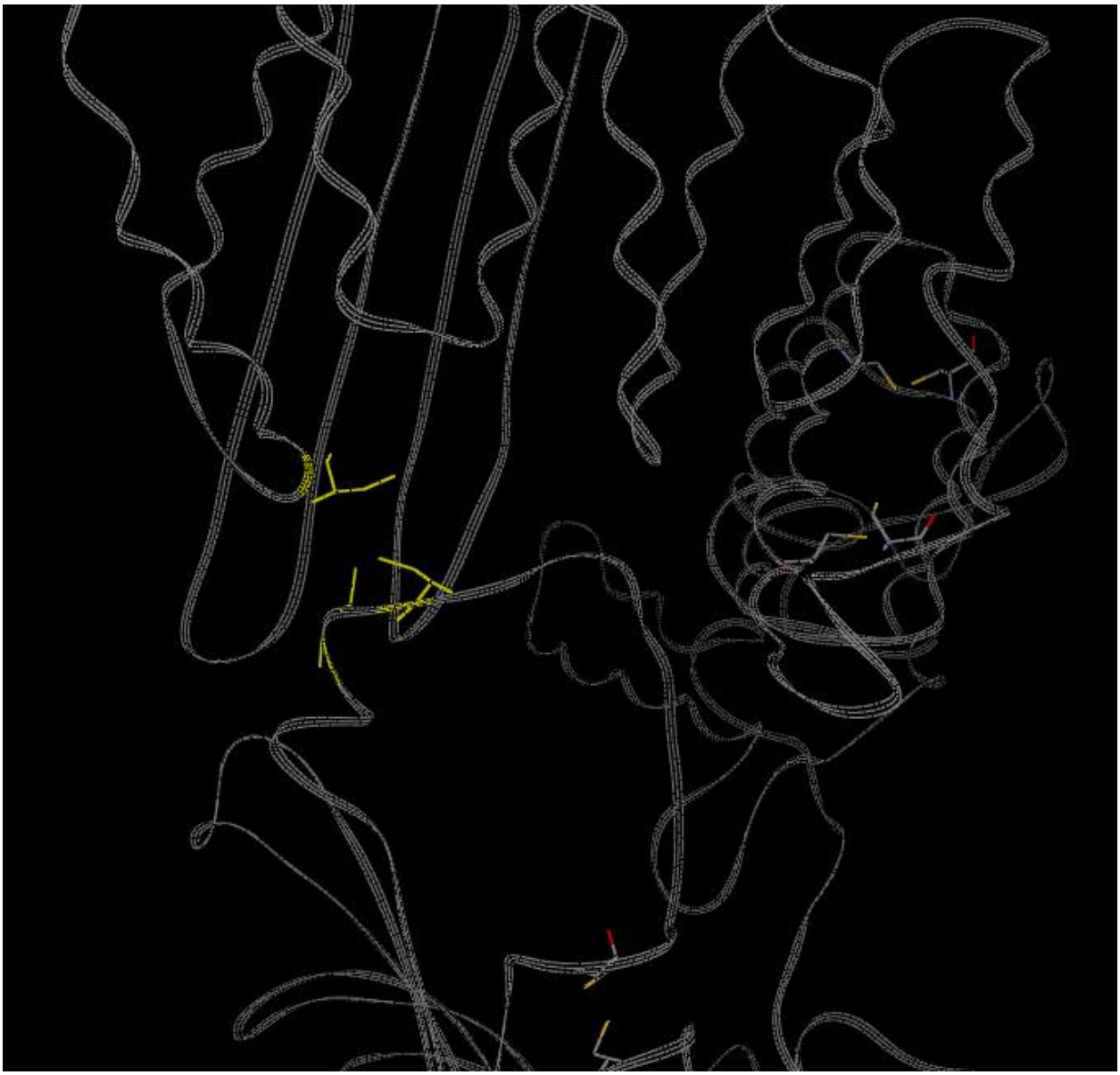


Fig. 23 Positions of C-186 and C-189 of the N-sheet, and C-616 of the helical domain of the modeled HeLp-B backbone. Cysteines of interest are highlighted in yellow. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0.

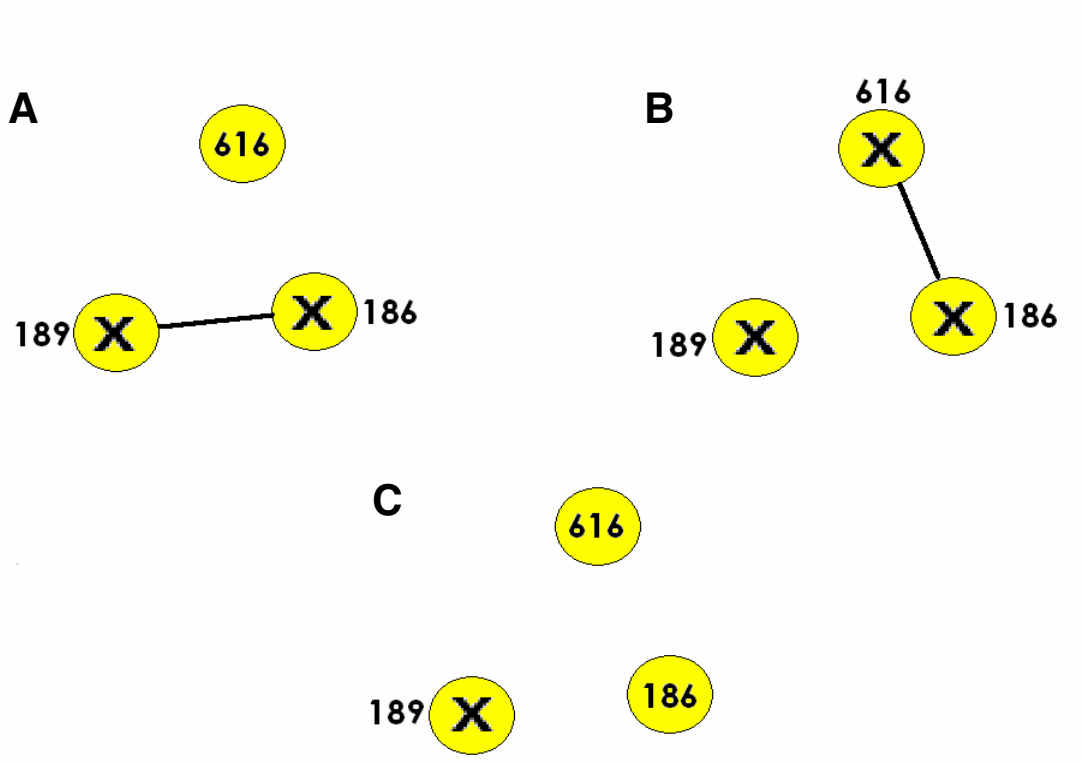


Fig. 24 Schematic diagram of the possibilities of disulfide formation for C-186, C-189, and C-616 in the rigid HeLp-B model. Solid circles represent the sulfhydryl groups of the cysteines. An "X" denotes that the particular cysteine is not available for linking with another HeLp-B subunit. A line between two cysteines represents an intramolecular disulfide bond between two cysteines. A cysteine without an "X" is available for disulfide bridge formation with another HeLp-B subunit.

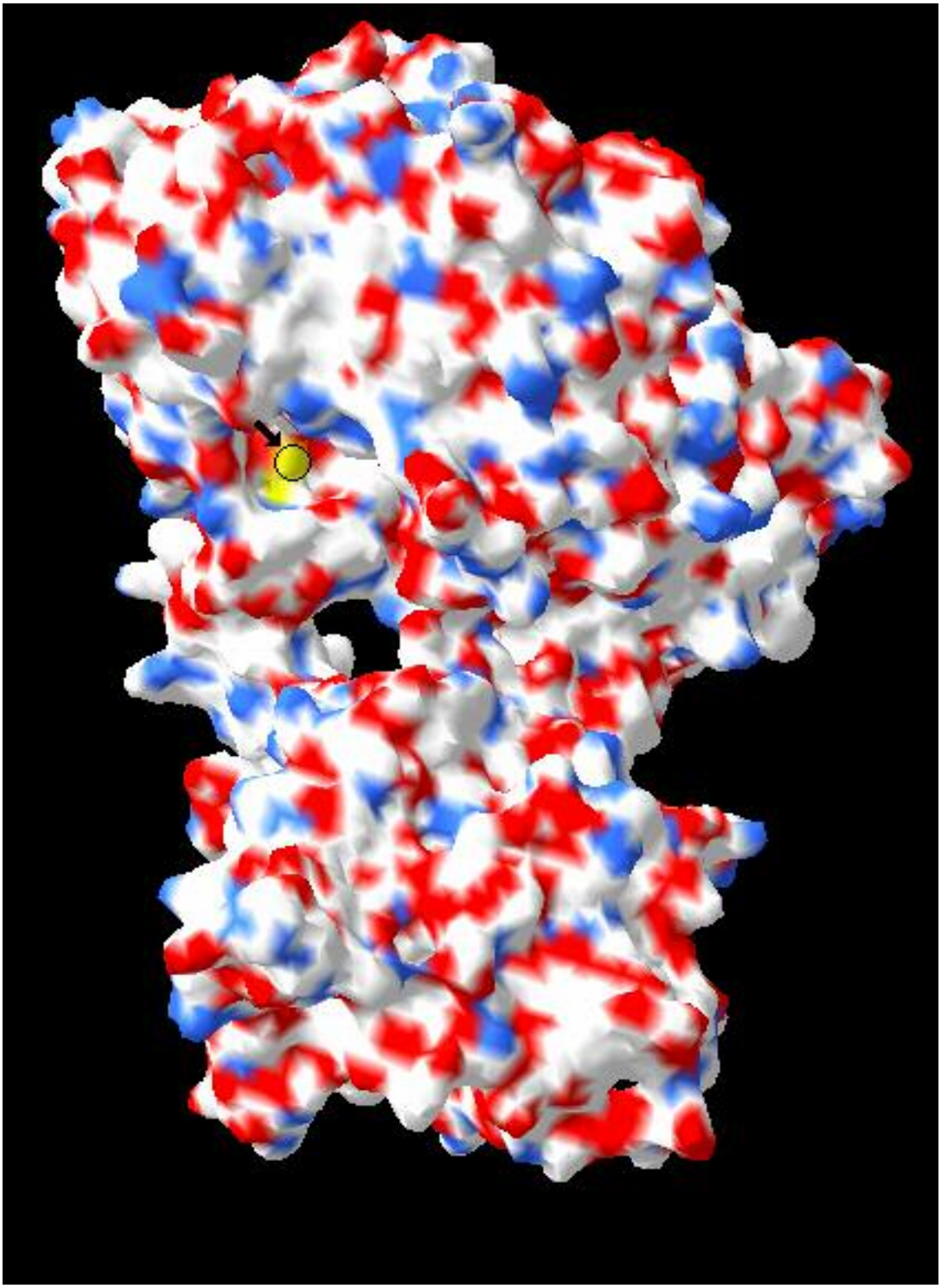


Fig. 25 Depiction of the calculated surface of the rigid HeLp-B model using the Conolly method and a probe radius of 1.5. The sulfhyryl of C-616, located in a cleft between the helical domain and the N-sheet, is circled and colored yellow. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0.

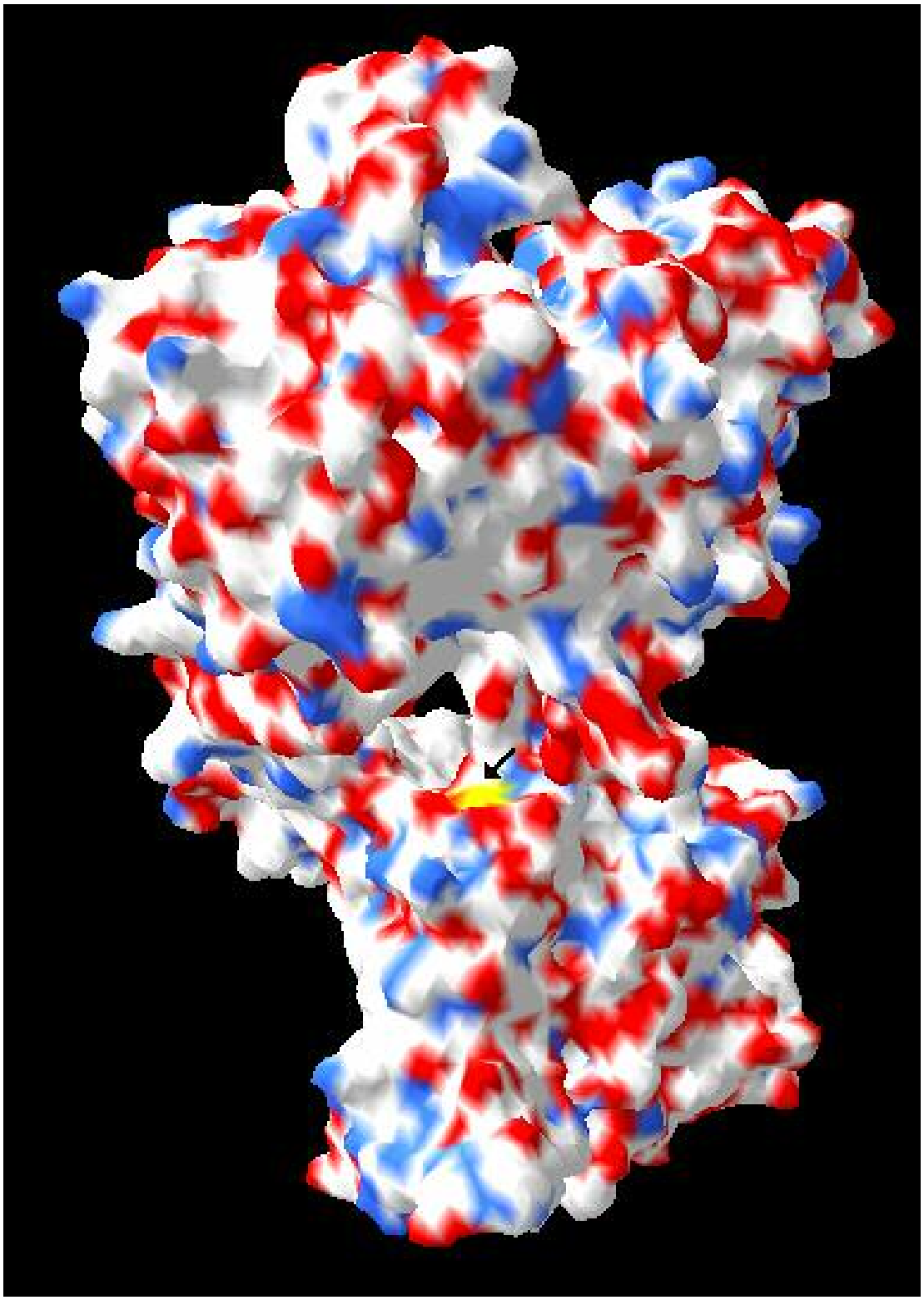
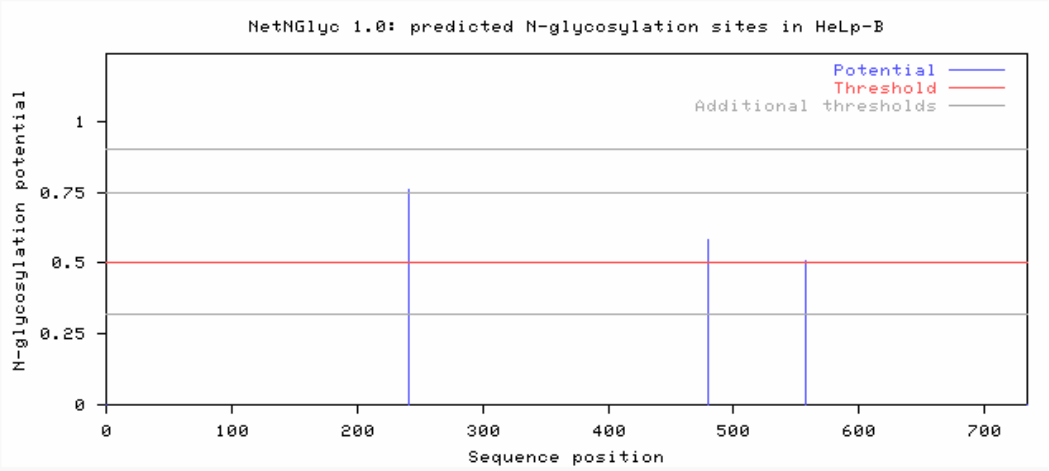


Fig. 26 Depiction of the calculated surface of the rigid HeLp-B model using the Conolly method and a probe radius of 1.5. The protein has been rotated from its perspective in the previous figure so that the N-sheet is further into the plain of page. The sulfhyryl of C-170, colored yellow, is located in the N-sheet and has been rotated to a position that makes it solvent accessible. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0.

A

(Threshold=0.5)

SeqName	Position	Potential	Jury agreement	N-Glyc result
HeLp-B	241 NRTL	0.7583	(9/9)	+++
HeLp-B	480 NYSV	0.5811	(8/9)	+
HeLp-B	558 NDSE	0.5090	(4/9)	+



B

(Threshold=0.5)

SeqName	Position	Potential	Jury agreement	N-Glyc result
HeLp-A	776 NNSL	0.5022	(5/9)	+

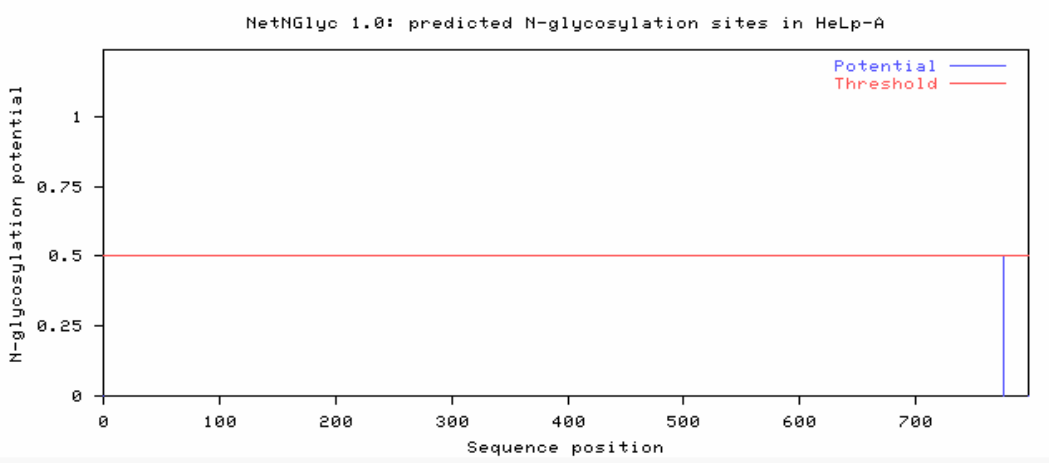


Fig. 27 The NetNGlyc 1.0 Server of the Center for Biological Sequence Analysis (http://www.cbs.dtu.dk/services/NetNGlyc/) (Gupta *et al.* 2004) prediction of glycosylation sites for (A) HeLp-B and (B) HeLp-A.

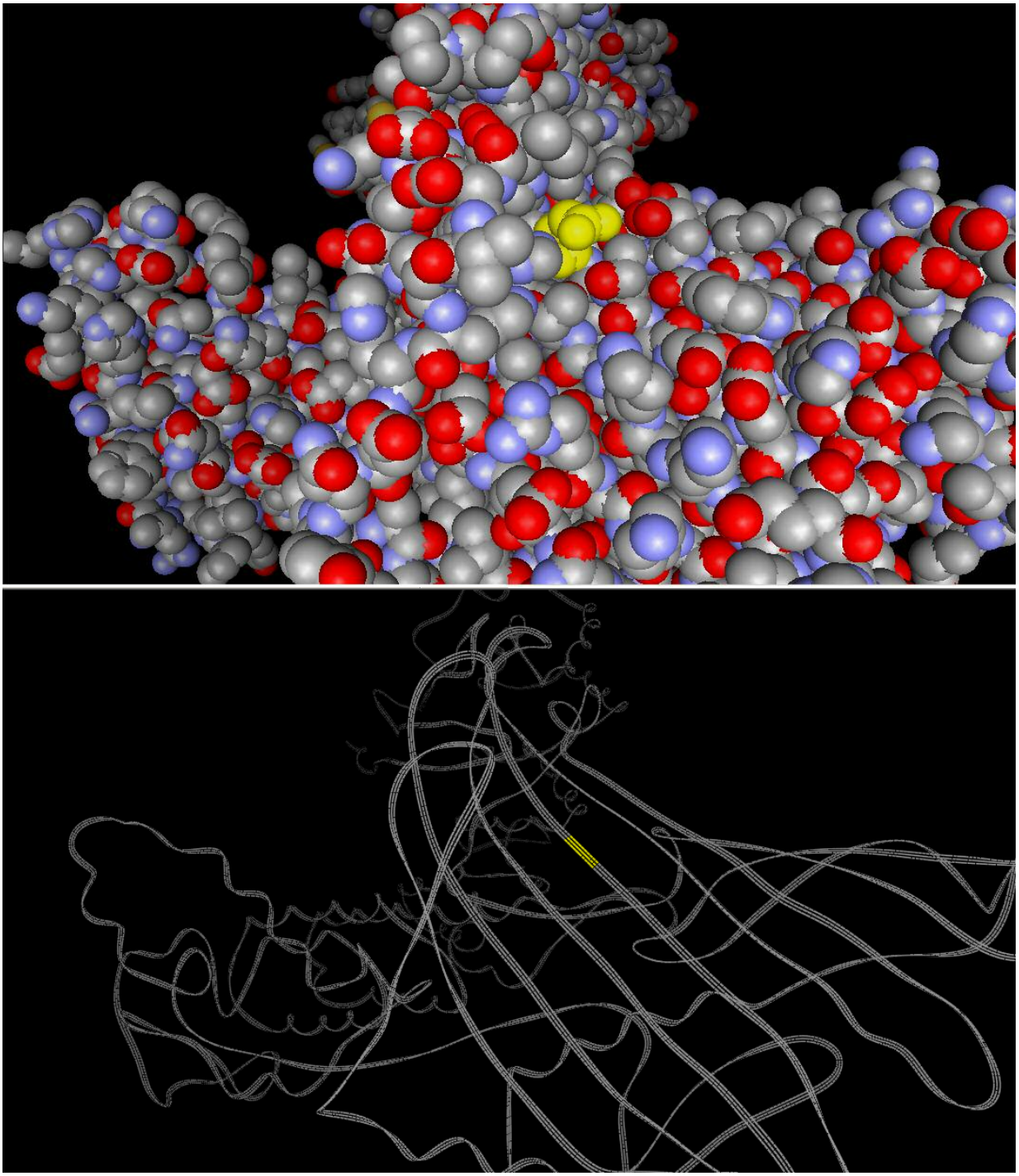


Fig. 28 Space-filling model and ribbon diagram of the predicted glycosylation site at asparagine-241 (colored yellow) on the N-sheet. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0.

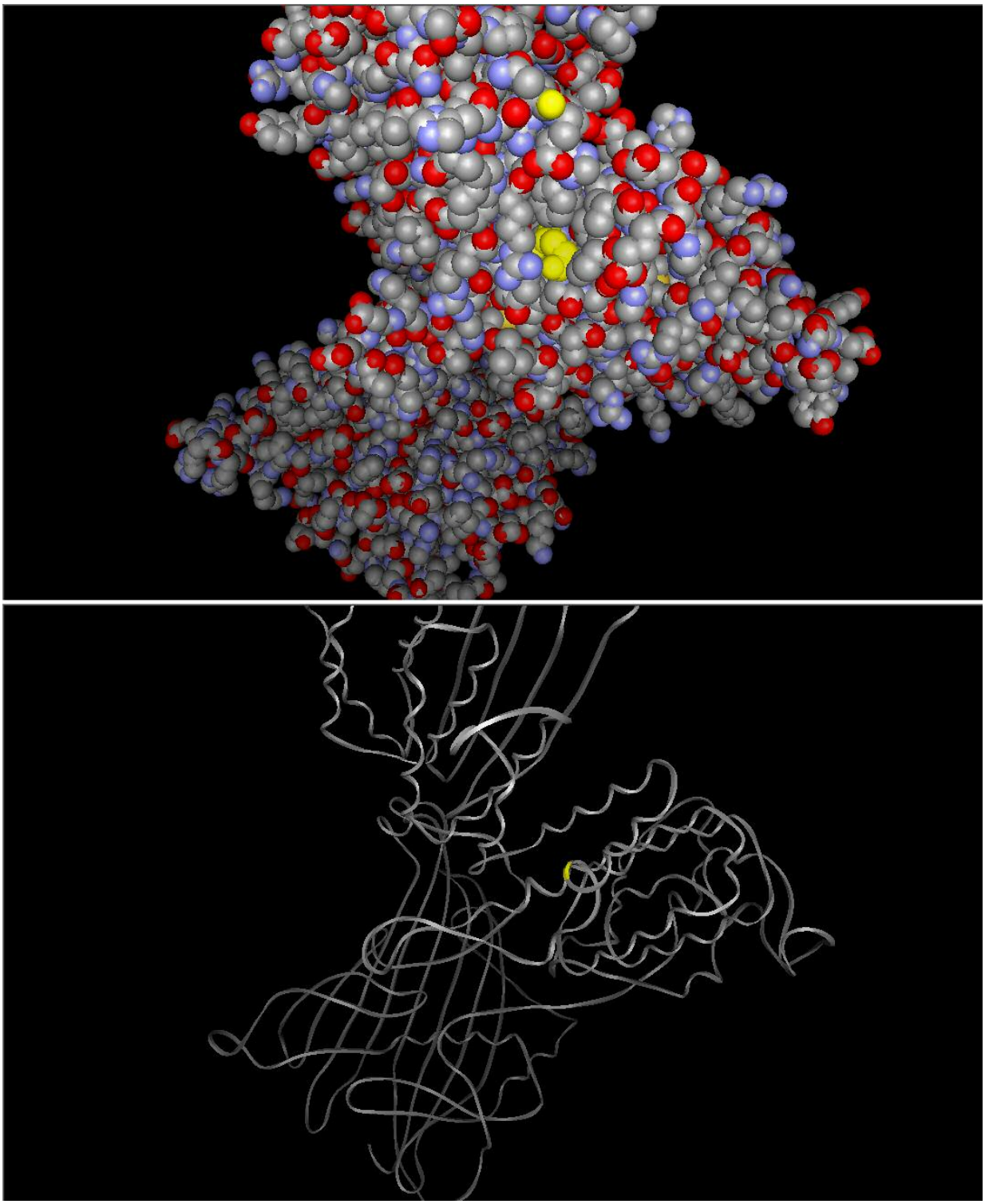


Fig. 29 Space-filling model and ribbon diagram of the predicted glycosylation site at asparagine-480 (colored yellow) on the helical domain. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0.

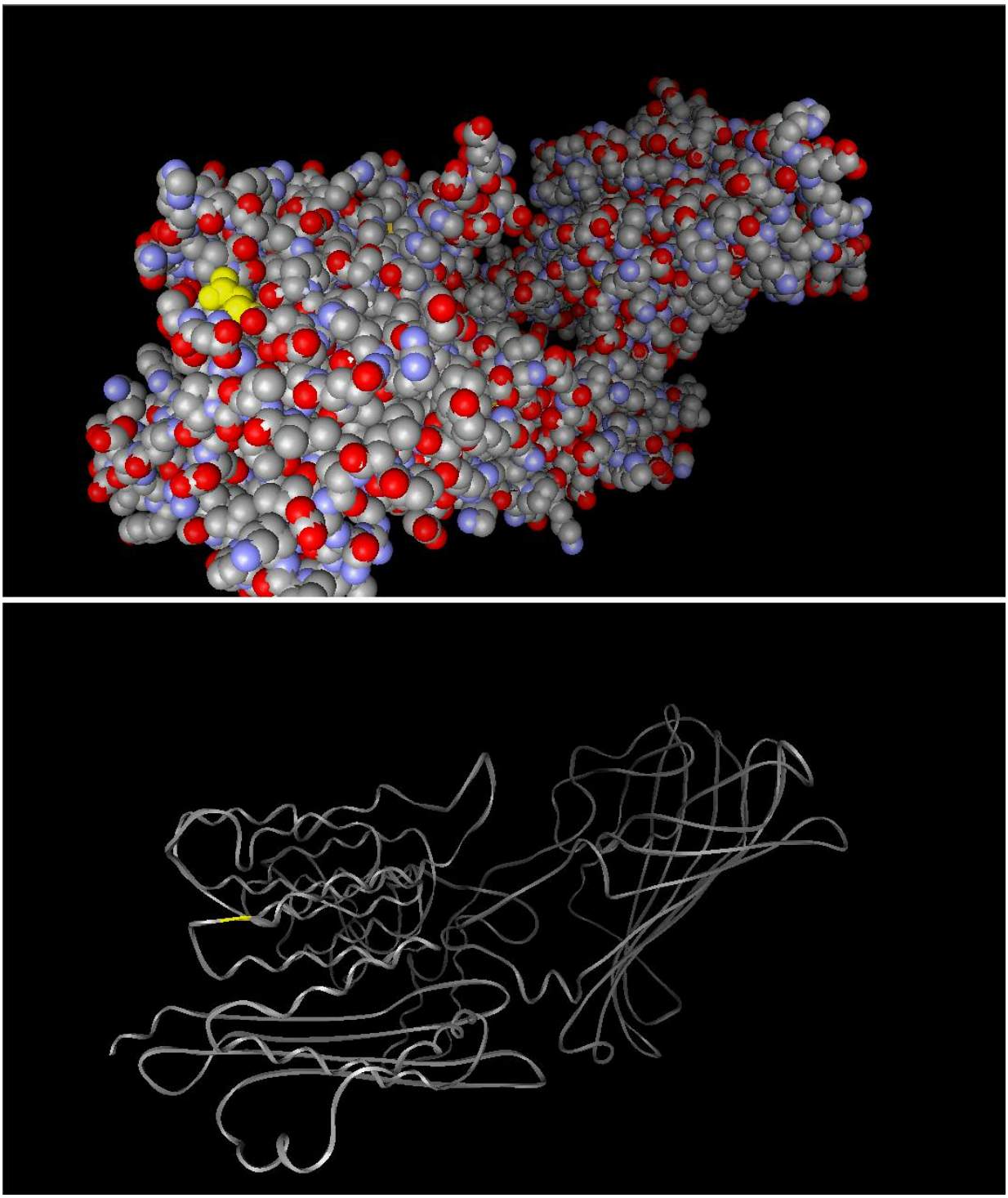


Fig. 30 Space-filling model and ribbon diagram of the predicted glycosylation site at asparagine-558 (colored yellow) on the helical domain. Figure created with 3D-Mol Viewer from Invitrogen's Vector NTI Advance 10.3.0

VITA

William Justin Cordill

Candidate for the Degree of

Master of Science

Thesis: CHARACTERIZATION OF HEME LIPOPROTEIN IN IXODID TICK SALIVA AND HEMOLYMPH

Major Field: Entomology

Biographical:

Education: Received a Bachelor of Science degree in Biochemistry and Molecular Biology from Oklahoma State University in May, 2005. Completed the requirements for a Master of Science degree in Entomology from Oklahoma State University in July, 2007.

Experience: Worked as a laboratory technician in a stored grain pests lab while earning my Bachelor's degree. The last semester of my Bachelor's degree I worked as a laboratory technician in an insect biochemistry laboratory. While earning my Master's degree I continued to work in the insect biochemistry laboratory while working at the OSU central tick rearing facility as a research assistant. My academic and professional background includes entomology, biochemistry, and molecular biology.