UNIVERSITY OF SOUTH BOHEMIA
Faculty of Science

Determination of amino acid sequence of hemelipoglycoprotein from tick *Dermacentor marginatus* by mass spectrometry

Bachelor thesis

David Pech

Supervisor:
prof. RNDr. Libor Grubhoffer, CSc.

Supervisors consultants:
Mgr. Ján Šterba and Mgr. Peter Koník

České Budějovice 2010

Declaration

I hereby declare and oath that I have elaborated the thesis on my own, only using the cited literature.

I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my dissertation thesis, in full / in shortened form resulting from deletion of indicated parts to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages.

In Budweis on May 24, 2010

.....
David Pech

Thanks

I would like to thank to Prof. Libor Grubhoffer for the opportunity to work on such interesting project and his guidance during the quest for the insight to the science. I would also like to thank to Mgr. Ján Štěrba. Without his patience during consultations and willingness to advise when necessary, I would not learn that much as I did during his guidance. My thanks also belong to Mgr. Peter Koník, who enlightened me with the knowledge of Mass Spectrometry and showed me practical approach to Mass Spectrometry. In addition, I would like to thank to Jarmila Dupejová, a person, without who there would not be any protein to study.

Contents

List of Figures

List of Tables

Annotation

English

The aim of the work was to get familiar with selected approaches that are used in mass spectrometry and to use them for determination of amino acid sequence of hemelipoglycoprotein (HLGP) from tick *Dermacentor marginatus* and to compare the obtained sequences with *D. variabilis* hemelipoglyco protein sequence from NCBI database. HLGP is a tick's lipoglycoprotein that binds and stores heme and together with vitellogenin (Vg), other tick's heme binding protein, plays important role in hematophagy (Donohue et al., 2008).

The work is part of a research project of the Laboratory of Molecular Biology of Vectors and Pathogens (a joint laboratory of the Institue of Parasitology, Biology Centre ASCR and Faculty of Science of the University of South Bohemia, České Budějovice), which is to identify, analyze and characterize FReD (fibrinogen-related domain) proteins from the ticks.

Czech

Cílem práce bylo seznámit se s vybranými postupy využívanými při hmotnostní spektrometrii a využití nabytých znalostí k určení aminokyselinové sekvence hemelipoglycoproteinu (HLGP) z klíštěte *Dermacentor marginatus* a porovnání získaných sekvencí s hemelipoglycoproteinovou sekvencí klítěte *D. variabilis* z NCBI databáze. HLGP je klíštěcí lipoglykoprotein, který váže a ukládá hem a společně s vitellogeninem (Vg), dalším klíštěcím proteinem vázajícím hem, hraje důležitou roli pro sání krve (Donohue et al., 2008).

Práce tématicky navazuje na projekt Laboratoře molekulární biologie vektorů patogenů (společná laboratoř Parazitologického ústavu Biologického centra Akademie Věd České republiky a Přírodovědecké fakulty Jihočeské univerzity v Českých Budějovicích), jejíž cílem je identifikace, analýza a charakterizace FreD (fibrinogen-related domain) proteinů u klíšťat.

Pech, D., 2010, Determination of amino acid sequence of hemelipoglyco protein from tick *Dermacentor marginatus* by mass spectrometry, Bachelor Thesis in English, 40 p., University of South Bohemia, České Budějovice, Czech Republic

Chapter 1

Introduction

1.1 Ticks

Ticks are important vectors of pathogens causing a variety of diseases in different host species, including humans. Pathogens transmitted by ticks include viruses, bacteriae and protozoans. This is enabled either because ticks tolerate the pathogens inside them, their immune system is not directed against them, or the pathogen does not harm the tick (Kovář, 2004). Tick injects saliva into the host which contains components that can modulate the local immune response of the host to enable the tick to feed. This phenomenon has been reported for most blood-feeding arthropodes, however the biggest number of examples has been recorded with ticks (Nuttall and Labuda, 2008). By injecting the saliva containing various pharmacologically active molecules secreted to the saliva, the tick higly modifies the site where it feeds. By doing so, the tick creates ecologically privileged niche that can be exploited by the pathogens and therefore it can facilitate the pathogen evasion from infected tick into the host tissue (Nuttall and Labuda, 2008). This phenomenon is in the literature referred to as Saliva-activated Transmission (SAT) (for review Nuttall and Labuda 2004).

1.1.1 What are ticks

Ticks are parasitic mites that attach firmly to the skin of the host. They feed on variety of hosts, such as humans and other mammals, birds, reptiles and rarely amphibians. Many of the ticks posess the ability to endure more than a year, while waiting for their next meal (Hillyard, 1996). Once on the host, the tick penetrates the skin with its mouthparts (creating only a small incision). Subsequent salivary secretion creates a lesion and allows the hypostome to be gradually worked in. In many species, salivary glands also secrete cement, which anchors the mouthparts in the host (Hillyard, 1996). Next, the tick infested with the blood increases its weight and uses the blood meal for subsequent egg production.

Classification

Ticks are mites (*Acarina*) and they are divided into 3 basic families:*Ixodidae* (hard-ticks), *Argasidae* (soft-ticks) and *Nutalliellidae*. The family *Ixodidae* contains 13 genera and 650 species. The family *Argasidae* has 5 genera and 170 species (Sonenshine, 1991). The family *Ixodidae* can be divided to 2 groups: Prostriata (genus *Ixodes*) and Metastriata (other genera such as *Dermacentor, Heamaphysalis, Hyaloma, Amplyoma*, etc). The families differ in onthogenesis, physiology, morphology or reproductory strategies. (Dupejová, 2008a)

Diseases

Ticks are blood-sucking arthropodes. which can be found in all terrestrial areas. They are the second most dangerous vectors of human and animal diseases after the mosquitos. They are main vectors of pathogens causing Lyme disease, Ehrlichiosis, Babesiosis, etc (Sonenshine, 1993). Only a minority of tick species are involved in the spread of medically and economically important diseases (Hillyard, 1996). Such species tend to be those which accept a wide range of hosts, including man. Those species are denoted as *vector competent* (Hillyard, 1996).

1.1.2 *Dermacentor marginatus*

Dermacentor marginatus, also known as the Ornate sheep tick, is present across Europe, Asia and North America. However, it is not present in the Britis Isles. The female is 4.6-5.4 mm big (engorged up to 15 mm). The male's size is 4.8-5.8 mm and the nymph is sized 1.4-1.8 mm (Hillyard, 1996).

All three developmental stages target different hosts. They target small and big mammals rarely attacking human. *Dermacentor marginatus* is the main vector of *Rickettsia slovaca* (Beati and Raoult, 1993; Řeháček et al., 1977).

1.2 Hemelipoglycoprotein

The hemelipoglycoproteins in *Chelicerata* binds and stores heme and thus it results in tranquilization ot the toxic effect of heme and utilizes it as a prosthetic molecule (Donohue et al., 2009). According to Donohue et al. (2009), the knowledge of heme-binding storage proteins is mainly derived from studies of vittellogenins (Vg) and ixodid hemolymph storage protein hemelipoglycoprotein (HLGP). However, the studies in this field are still in the beginnings and HLGP have to be studied further. The HLGP is closely related to vitellogenin in terms of primary structure, but differs in post-translational processing (Donohue et al., 2008). Hemelipoglycoprotein consists of 2 subunits. Together they contain three motifs:

- a lipoprotein N-terminal domain (common attribute of lipid, carbohydrate and metal binding proteins)

- a domain of unknown function characteristic for proteins with several large open beta sheets
- a von Willebrand factor type D domain near the C-terminus

(Donohue et al., 2008)

1.2.1 Function

According to Donohue et al. (2008) the presence of the lipoprotein N-terminal domain supports the putative storage function of this protein. The HLGP together with Vg seems to be critical component for sequestration and utilization of host heme in the absence of *de novo* synthesis of heme in the *Ixodidae* and for utilization of heme during extended periods of nonfeeding. The expression in female adults is initiated by attachement to the host and blood feeding (Donohue et al., 2008). The HLGP was detected in the fat body, salivary glands, ovary and midgut. However after 6 days of feeding the levels were higher in the fat body and salivary glands, than in the ovary and midgut (Donohue et al., 2008; Guddera et al., 2002).

The HLGP binds heme for utilization by the tick and also provides a sequestration mechanism to prevent the formation of potentially damaging radical species (Maya-Monteiro et al., 2004). The sequestration is carried out in the hemolymph in a 2:1 heme to HLGP ratio (Maya-Monteiro et al., 2000).

HLGP together with Vg appears to have critical role also in the evolution of hematophagy (Donohue et al., 2008).

1.3 Mass spectrometry

Mass spectrometry is an analytical technique that allows identification of the substance based on its mass. The measurement actually estimates the mass-to-charge (m/z) ratios, because the substance is ionized during the analysis. Mass spectrometry can be successfully used for structure elucidation of molecules such as peptides and proteins, based on the masses of their fragments.

1.3.1 Mass spectrometer

Typical mass spectrometer consists of 3 main parts. They are the ion source, mass analyzer and detector. These are further described in more details.

Ion source

Ion source is a device that creates gas phase ions. The atoms and molecules in the sample are transferred into the gas phase and then ionized concurrently (eg. Electro Spray Ionization ESI) or through separate processes (eg. glow discharge). So called soft ion sources can produce intact ions of large fragile molecules, such as in ESI.

ESI stands for Electro Spray Ionization. The concept was introduced by Dole and coworkers in 1968 and was coupled to mass spectrometry by Yamashita and Fenn in 1984. The sample is dissolved in a polar volatile solvent and transported through a needle placed at high positive or negative potential (relative to a nozzle surface). This potential between the needle and the nozzle forces the fluid to create a Taylor cone (Westman-Brinkmalm and Brinkmalm, 2009a).

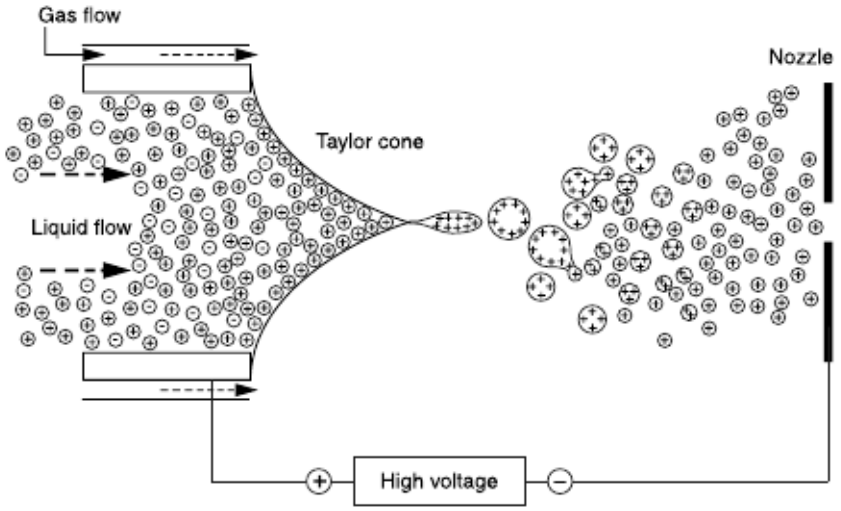

Figure 1.1: Schematic drawing of ESI source with Taylor cone depicted (Westman-Brinkmalm and Brinkmalm, 2009a)

The droplets shrink during the evaporation process, assisted by a flow of hot N_2 gas that passes across the front of the ionization source. The droplets not only shrink, but are also fragmented by increasing concentration of the positive charge on the shrinking droplet resulting in its disruption. Ions are formed at atmospheric pressure and pass through the cone shaped inlet into the intermediate vacuum region and from there they pass through a small aperture into the high vacuum of the MS analyzer (Westman-Brinkmalm and Brinkmalm, 2009a).

The sample is dissolved to a suitable concentration in a mixture of H_2O/organic solvent (methanol, isopropanol or acetonitrille). A small amount of formic or acetic acid is added to aid protonation of the analyte molecules for positive ionization mode and to maintain uniform condition durin LC separation. For negative ionization mode ammonia solution or volatile amine are added to aid deprotonation of the analyte (Westman-Brinkmalm and Brinkmalm, 2009a).

Collision induced/activated dissociation

During collision induced/activated dissociation (CID/CAD) the precursor ions collide with atoms or molecules of inert gas resulting in fragmentation of the precursor ion. In the collision, part of the kinetic energy is converted into vibrational/rotational energy of the parent ion. If the internal energy gain is high enough, the precursor ion will fragment (fragmentation - see section 1.3.2) fast enough to be observed in the mass spectrometer (Westman-Brinkmalm and Brinkmalm, 2009b). There are 2 types of CID used, depending on the type of mass analyzer. The high-energy CID is typical for

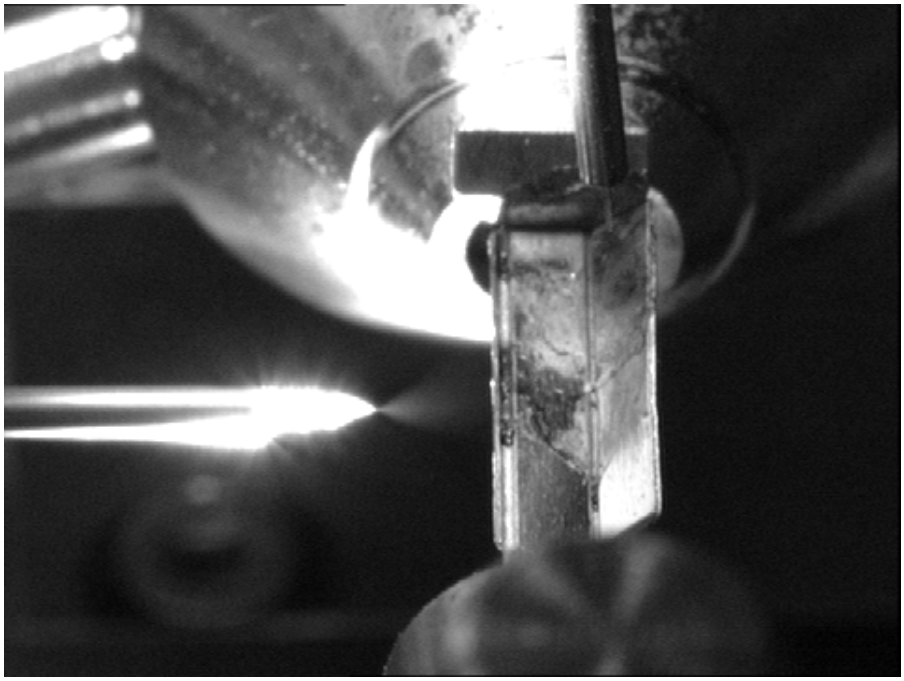

Figure 1.2: Photo of ESI ion source

sector and TOF instruments, whereas the low-energy CID is typically used with quadrupole and ion trap instruments.

Mass analyzer

Mass analyzer is a device that separates the generated ions according to their masses. The principle is based on electromagnetism. For the purpose of this work, 2 types were used: Quadrupole and time-of-flight analyzers.

Quadrupole is a mass filter device that can be set to allow to fly through only ions with limited m/z range. The ions with either higher or lower m/z will not escape the filter. The mass spectrum can be obtained by scanning through the whole m/z range of interest and detection of ions that pass at each m/z.

The principle was first published in 1953 by Paul and Steinwedel. The device consists of 4 symmetrically arranged parallel rods (see Figure: 1.3) with combination of direct-current (DC) and radio-frequency (RF) potentials. Opposite rods are connected in pairs. Those 2 pairs have potential of same magnitude, but opposite polarity. Accelerated ions enter region between the rods and only the ions with certain narrow m/z ratio will have stable trajectories at given DC and RF potentials and RF frequency. Other ions with m/z values outside of the range will collide with the rods, which causes loss of charge and therefore prevents the ions to reach the detector.

Heavy (more inert) ions respond to the DC component, whereas the lighter (quicker) ions respond to the altering RF component. One pair of the rods serves as a high-pass filter and second pair serves as a low-pass filter. (Westman-Brinkmalm and Brinkmalm, 2009a)

- **High-pass filter**
  The trajectory of the heavy ions is directed to the middle between the rods. The low-mass ions respond faster and are attracted to the rod, when once every cycle, the net force of the DC and RF component becomes attractive.

- **Low-pass filter**
  Heavy ions experience attraction and collide with the rod. Low-mass ions are also attracted, but once every cycle, they experience repulsion to the middle section between the rods.

If the alternating field between the quadrupole rods is tweaked appropriately, the ions with desired m/z range will have stable trajectories.

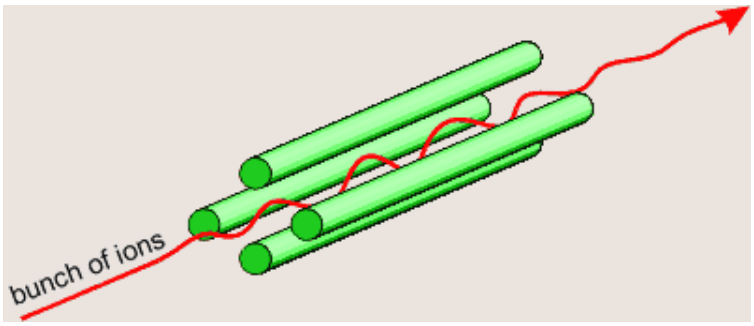

Figure 1.3: Schematic drawing of quadrupole mass analyser used in mass spectrometers, picture reprinted from Wikipedia under the terms of the GNU Free Documentation License, Version 1.2 (Wikipedia, 2010b)

The time-of-flight (TOF) mass analyzer was first mentioned in 1946 by Stephens (Westman-Brinkmalm and Brinkmalm, 2009a). The used negative or positive high potential accelerates the ions created in the ion source. The ions ideally get the same kinetic energy, but since they differ in their respective m/z ratios, they have different mass dependent velocities and thus the time they spend on they way to the detector varies. The time between the start of the signal and the pulse generated when the ion hits the detector can be used to estimate the m/z ratio of the respective ion.

$$t_{TOF} = \frac{L}{\upsilon} = L\sqrt{\frac{m}{2qU_a}} \propto \sqrt{m/z}$$

L is the length of the field-free region, υ is the ion velocity after acceleration, q is the charge of the ion, m is the mass of the ion, z is the charge state and U_a is the accelerating electric potential difference (Westman-Brinkmalm and Brinkmalm, 2009a).

The resolution of this mass analyzer is limited by the initial velocity spread of the ions. Therefore it is necessary to compensate for the initial velocity distribution. The compensation is done by time focusing devices. There are 2 types of time focusing devices: the Electrostatic ion reflector and delayed extraction (time-lag focusing).

Electrostatic ion reflector was first introduced in 1973 by Mamyrin et al. The ions are reflected by a higher potential than the acceleration potential. After this, they hit the detector which is placed off axis.

The reflector increases the flypath of the ion. The slightly faster ions penetrate deeper into the electric field gaining longer flypath than the slower ones. This compensates the initial velocity spread. By optimizing the ratio of acceleration and reflection voltages, the ions with the same m/z, but different initial velocities will arrive to the detector at (almost) same time (Westman-Brinkmalm and Brinkmalm, 2009a).

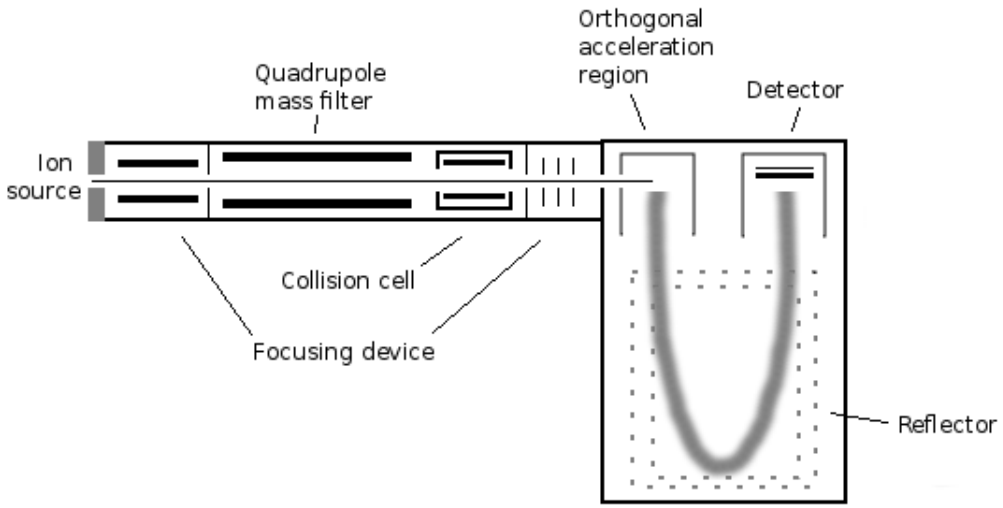

Figure 1.4: Schematic drawing of quadrupole time-of-flight (qTOF) mass spectrometer, drawn in Gimp

Detector

The purpose of a detector in mass spectrometry is co convert energy of incoming particles into a current signal that can be registered by electronic devices and transferred to a computer. The impact of an incoming particle causes an emission of secondary particles (electrons). The number of secondary particles depends on velocity and/or energy of the incoming ion.

If all ions are accelerated to the same kinetic energy (eg. in TOF), then the sensitivity for high-mass (slow) ions is lower than for low-mass (quick) ions. To increase the sensitivity, post-acceleration of the ions should be applied before the detector (Westman-Brinkmalm and Brinkmalm, 2009a).

Prerequisities for a mass spectrometry detector:

- High efficiency for converting energy of the ion to electrons
- linear response
- low thermionic emission (low background noise)

- short recovery time
- minimal variations in transit state

Electron multiplier amplifies a weak current of incoming particle by using a series of secondary electrodes or dynodes to produce higher current at the anode. The dynode often consists of an alloy of alkali/alkali earth metals with a more noble metal.

The Microchannel plate or Multichannel plate detector (MCP) consists of paralell array of continuous dynode channel electron multiplier. The inner surface is treated with a semiconducting material as secondary electron emitter. The flat ends of the channels are coated with a metallic alloy to enable application of potential difference over the MCP. This potential accelerates the particles inside of the dynode channel towards the output end.(See Figure: 1.5). When an ion impacts the wall of the channel, an avalanche of electrons is created. Each channel acts as continuous dynode and can generate up to 10^4 electrons per impact. (Westman-Brinkmalm and Brinkmalm, 2009a) Usually 2 MCP detectors are used in tandem setup to increase the gain of the detector (See Figure: 1.6).

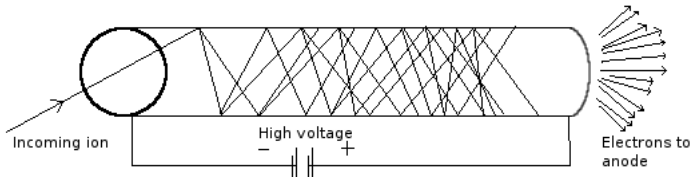

Figure 1.5: Single MCP channel schematics, drawn in Gimp

Major drawback of MCP detector is a saturation effect. The electron cascade in one channel drains other neighbouring channels for several microseconds. This produces non-linearities for count rates above few thousand incident particles per second (Westman-Brinkmalm and Brinkmalm, 2009a).

1.3.2 Sequencing of proteins and peptides

Tandem mass spectrometry

Tandem mass spectrometry (MS/MS) is an analytical technique which is used to get structural information about the sample molecule by using multiple stages of mass selection and mass separation (Westman-Brinkmalm and Brinkmalm, 2009b).

Concept For peptide sequencing positive mode is used. Proteins when ionized usually produce pseudomolecular ions (Noga et al., 2009), which means that molecules acquire charge-carrying protons. This proton must be taken into consideration when predicting the correct m/z ratio. For example, a protein with a molecular weight (MW) 3000 Da, singly charged, will be detected at m/z 3001 (considering proton equal as 1). When doubly charged, the same protein will be detected at m/z 1501.

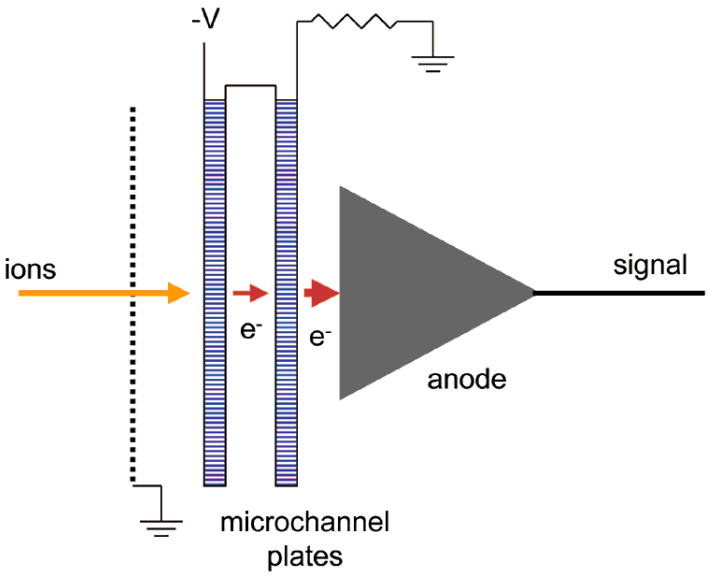

Figure 1.6: Schematic drawing of microchannel plate detector, picture reprinted from Wikipedia under the terms of the GNU Free Documentation License, Version 1.2, (Wikipedia, 2010a)

$MW = 3000$
$m/z = (3000 + 1)/1 = 3001 (z = +1)$
$m/z = (3000 + 2)/2 = 1501 (z = +2)$
$m/z = (3000 + 4)/4 = 751 (z = +4)$
etc.

The charge state of proteins generally depends on (Noga et al., 2009):

- Ionization method - ESI tends to create multiply charged pseudomolecular ions compared to Matrix Assisted Lased Desorbtion/Ionization (MALDI), where multiple ionization is not so frequent
- Peptide lenght - longer chains have more groups, where additional protons can attach, therefore they have bigger chance to generate multiply charged ions
- Amino acid sequence - some amino acids tend to be more ionized than others (eq. arginine or lysine)
- Instrument settings, composition or pH of solvent

Classical approach in protein sequencing The acquisition of protein structural information is based on the fact, that MS/MS spectrum of fragmented peptide will yield a ladder of peaks, where distances between peaks reflect the sequence of amino acid residues. However, the bigger the peptide is, the harder is the correct interpretation of the spectra. The approximate upper limit is 2 to 4 kDa (Noga et al., 2009). Therefore the first step in the method is to fragment the protein to smaller peptides in a predictable way. Second step is acquisition of a parent spectrum (first stage of tandem mass

spectrometry) of the protein digest, isolation of a precursor (one digested fragment from the mixture), this is where the quadrupole mass filter is employed. Third step is fragmentation of the precursor and acquisition of a fragmentation spectrum (second stage of tandem mass spectrometry). This is repeated for each fragment (precursor) from the digest mixture. In the ideal case. the amino acid sequence of the peptide can be acquired from the fragmentation spectra.

Peptide fragmentation Peptide sequencing in mass spectrometry is not easy because various covalent bonds can be disrupted during fragmentation in the collision cell (see section 1.3.1). The bond breaking can occur both on the backbone and side chain groups. Therefore to precisely describe what happens, special nomenclature was devised.

Roepstorff's nomenclature introduced by Roepstorff and Fohlman in 1984 and modified by Johnson et al. in 1987 describes the fragments by lower-case letter with additional indexes (see Figure: 1.7). The ions derived from the C-terminus of the original peptide are noted x,y,z and those from N-terminus are noted a,b,c. Numerical subscripts inform about the number of amino acids residues present at given ion.

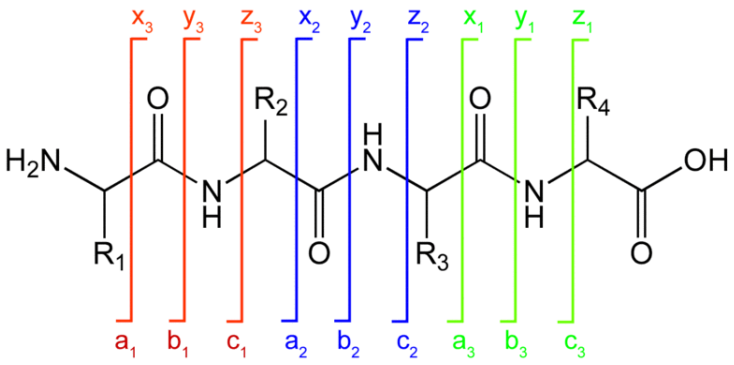

Figure 1.7: Scheme of peptide fragmentation according to Roepstorff and Fohlman, picture reprinted from Wikipedia under the terms of the GNU Free Documentation License, Version 1.2, (Wikipedia, 2010c)

QTOF MS/MS This mass spectrometer couples together 2 different types of mass analyzers. In MS mode the quadrupole allows transmission of ions with wide range of m/z ratios. In MS/MS mode, the quadrupole serves as a mass filter for the precursor ion of interest. The fragmentation of the precursor ion occurs in the collision cell through low-energy CID. The fragment mass spectrum is acquired by the TOF (Westman-Brinkmalm and Brinkmalm, 2009b).

De novo sequencing

The process of de novo sequencing of proteins is in fact the same as for tandem mass spectrometry of proteins (see 1.3.2). The only difference lays in fact that the amino acid sequence obtained is not compared to a protein database, but the spectrum is used to construct a new one. Knowing the

specificity of the enzyme used for digestion, it is possible to predict the C-terminal amino acid of the peptide fragment

Chapter 2

Methods

The glycoprotein was obtained from Jarmila Dupejová. It was prepared according to the experiments of her bachelor thesis (Dupejová, 2008b).

2.1 Materials and solutions

- 1% formic acid (Fluka)
- ammonium bicarbonate buffer 0.1 M (Sigma)
- acetonitrille LC-MS grade (Sigma-Aldrich) 100% and 50% mixture with 1% formic acid
- Trypsin Proteomics Grade (Sigma-Aldrich)
- V8 (Sigma-Aldrich)
- MiliQ H_2O

2.2 Software

- mMass 3.3.0 (Strohalm et al., 2010)
- BioEdit 7.0.9.0 (Hall, 1999)
- ProteinLynx Global Server 2.3 (Waters)
- LaTeXtypesetting software
- Vim (Vi improved text editor) 7.2
- OpenOffice 3.1
- Gimp (GNU image manipulation program) version 2.6

2.3 Protein Digestion

The protein was diluted into the final concentration of 10 μg/μl and mixed with proteolytic enzyme (either trypsin or V8), to the final concentration of the 0.2 μg/μl of the enzyme and incubated 10-16 hours at 37°C. The protein and the enzyme were diluted with ammonium bicarbonate buffer (ABC).

2.4 Protein digest treatment before MS

The protein digest was purified through C-18 ZipTip (Millipore) filters, so that the mixture contained only the peptide fragments.

Sample preparation steps with ZipTip C18:

- wash 3x with 10 μl of 100% acetonitrille
- wash 3x with 1% formic acid
- aspirate 15-20x 5 μl of the solution with digested protein
- wash 7x with 1% formic acid
- elute the peptides 10x with 20 μl of solution of acetonitrille/formic acid (50%/1%), both washes are collected into a sample vial and let evaporate for 1 hour
- clean the ZipTip 3x with 10 μl of 100% acetonitrille
- After eveporation of 50% acetonitrille, add 10 μl of H_2O into the sample to fill it the original volume

2.5 MS settings

HPLC: Nano Acquity UPLC (Waters, UK)

Trapping: Symmetry C18 Trapping column (180μm i.d. x 20mm length, particle size 5μm, reverse phase; Waters, UK)

Analytical column: BEH300 C18 analytical column (75μm i.d. x 150mm length, particle size 1.7μm, reverse phase; Waters, UK)

Solutions: A- Milli Q water + 0.1% formic acid, B - acetonitrille LC-MS grade + 0.1% FA

dda: The spectrometer performs the measurement in MS mode, with collision energy 5V. When an ion is detected, the machine switches to MS/MS mode, with collision energy 15-35V.

1 min trapping	99% A		15μL/min
beginning	97% A	3% B	400nl/min
30min	60% A	40% B	-ll-
32min	15%A	85% B	-ll-
37min	15%A	85% B	-ll-
40min	97%A	3% B	-ll-
end			

Table 2.1: Gradient (default)

LockMass calibration: During the measurement, internal standard is measured in parallel to ensure consistency during acquisition of the data. GluFib in 50% acetonitrille + 0.1%formic acid was used as the standard with settings to 2 masses, 785.8426 for 2x charge a 684.3469 for 1x charge with tolerance for both 0.3Da and calibration error less than 3ppm

Resolution 10 000

Software: PLGS 2.3 (Protein Lynx Global Server 2.3) for deconvolution, database search and *de novo* sequencing. Database Uniprot species specific (*Ixodidae*) was used. The analysis took into account 1 missed-cleavage site.

With regard to similarities between *Dermacentor marginatus* and *D. variabilis*, only the sequences, which corresponded to hemelipoglycoprotein sequence of *D. variabilis* were selected.

Tolerance: peptide accuracy 50ppm and MS/MS fragment mass accuracy 0.2 Da.

Chapter 3

Results

Hemelipoglycoprotein sequence from *Dermacentor variabilis* was chosen as a reference sequence for sequencing of hemelipoglycoprotein from *D. marginatus*.

The digestion was performed 3 times. Each digest was analyzed by the Mass Spectrometer 3 or 4 times.

The de novo sequencing of hemelipoglycoprotein (HLGP) from *D. marginatus* was carried out with resulting coverage of 8.15% of total amino acids residues (126 amino acids residues from total number of 1546) for the first subunit and 10.66% of total amino acids residues (130 amino acids residues from total number of 1492) for HLGP 2 subunit. The coverage is the ratio of successfully identified amino acids to complete sequence of the protein. This ratio is expressed in percentage.

The protein was digested by trypsin and V8 enzymes and spectra from tandem mass spectrometry were recorded. The procesing of the data and sequence determination was performed using Waters PLGS software (see section 2.2).

The fragments were grouped together according to their molecular masses (the molecular mass of the fragment was determined by PLGS 2.3 - see 2.2) and one representative sequence was chosen from each group according to the ladder score, which depicts the quality of the obtained sequence from the fragmentation spectra.

The chosen peptide fragment sequences were BLASTed (BLAST as integrated feature in BioEdit - see 2.2) with UniProt databases for *Ixodidae* and *Rhiphicephalinae*. The database search found similarities to hemelipoglycoprotein, hemelipoglycoprotein 2 and vitellogenin from *D. variabilis*.

In total, results of trypsin and V8 digests contained 19 fragments with sequence matching HLGP and 13 fragments with sequence matching vitellogenin (see tab:3.1), both from the tick *D. variabilis*.

Trypsin digest showed better results than V8, because from the total number of matches, V8 digest contained only 8 sequence fragments of HLGP and no vitellogenin matching sequence fragments, compared to trypsin digest with 11 sequence fragments of HLGP and 13 of vitellogenin.

The algorithm used for pairwise alignment of the fragments was ClustalW (with default parameters) embedded as a standard feature in BioEdit software. Those sequences that occured during the experimets multiple times were treated as single sequences. Merging all the pairwise aligned

	hemelipoglyco proteins	vitellogenin
trypsin	11	13
V8	8	none

Table 3.1: Table of number of matched sequence fragments per digest.

sequences yields the final result as seen in fig: 3.1, fig: 3.2 and fig: 3.3.

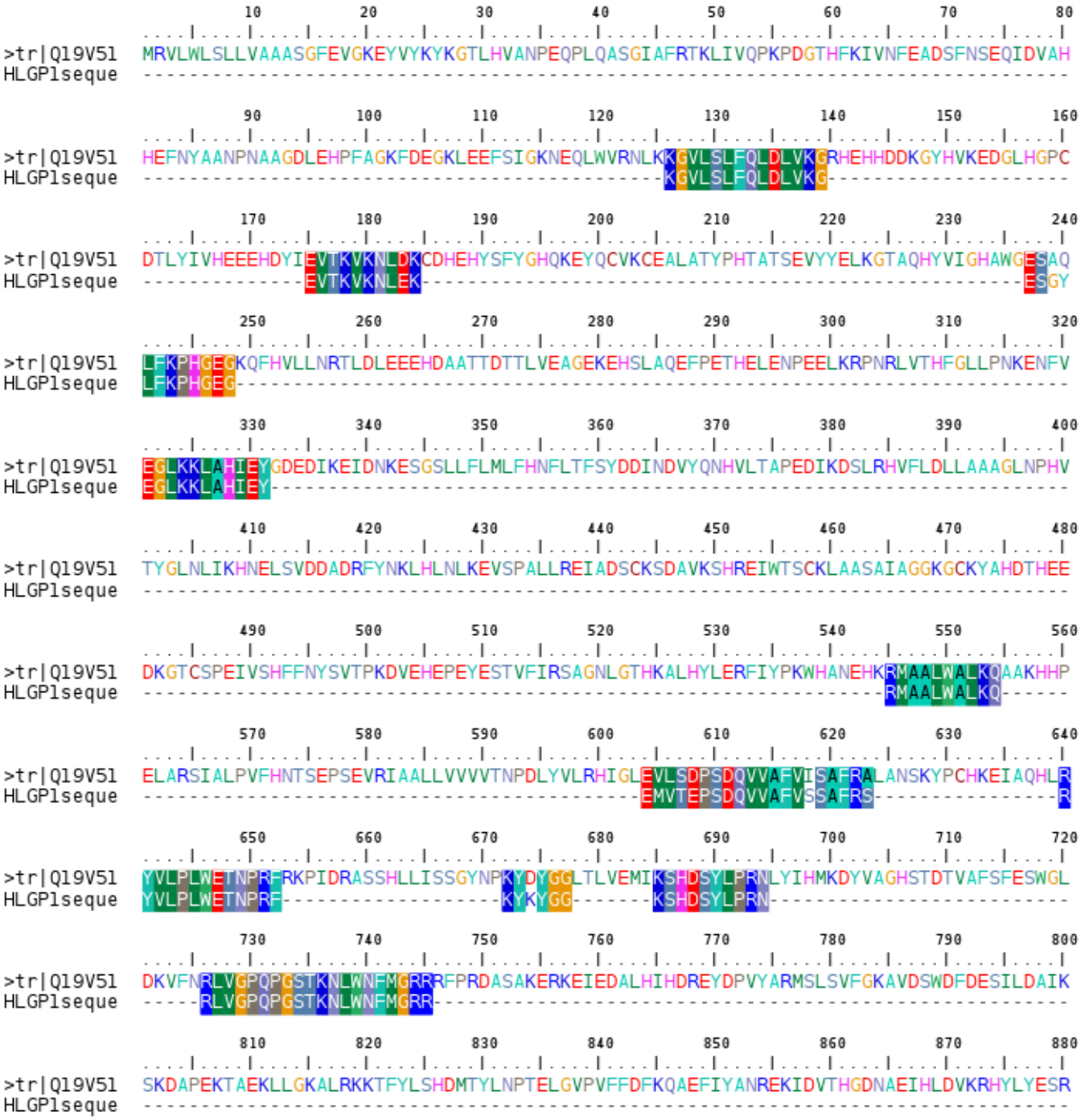

Figure 3.1: Comparison of *D. variabilis* hemelipoglycoprotein 1 subunit with de novo sequenced protein from *Dermacentor marginatus*.

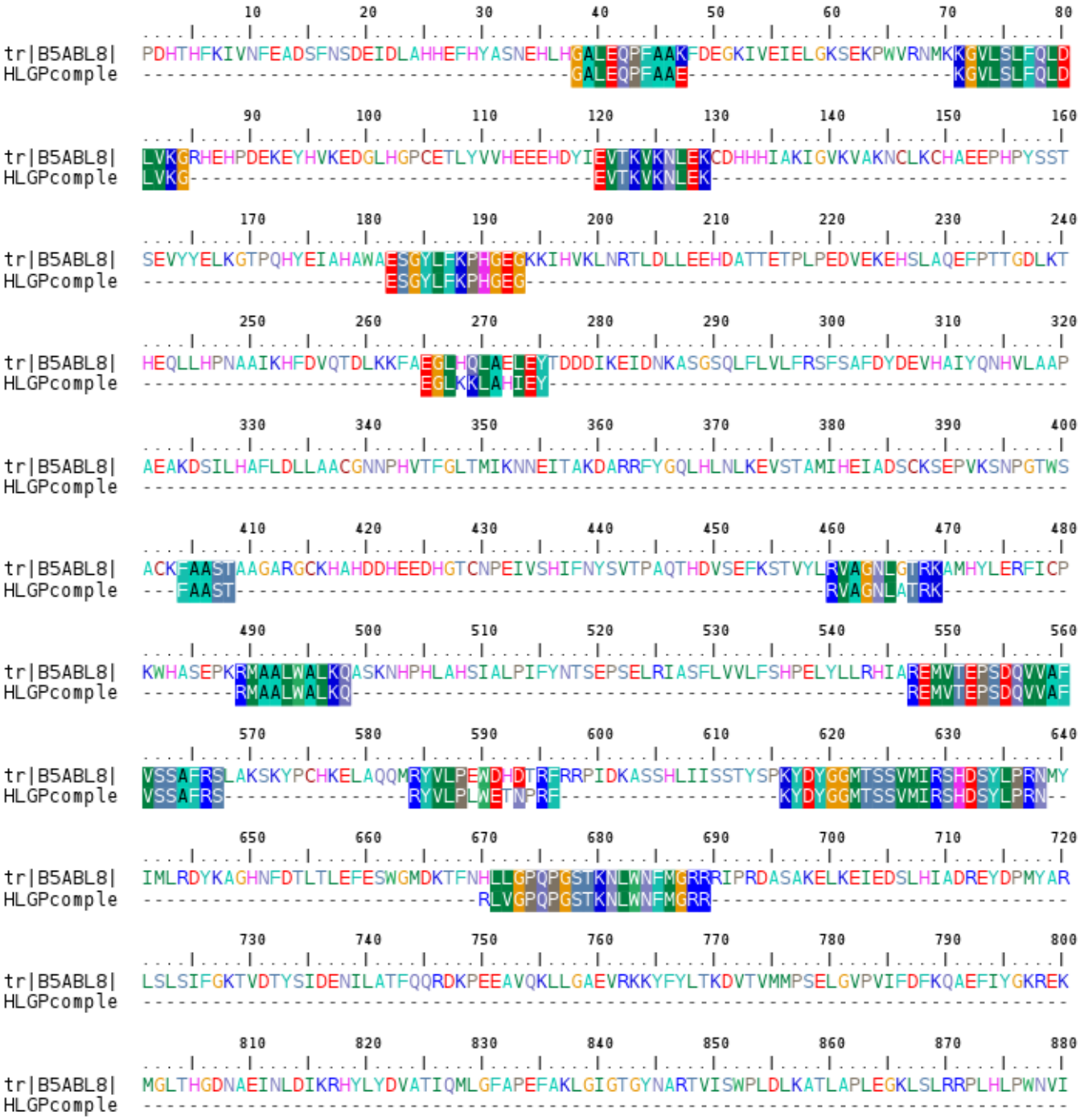

Figure 3.2: Comparison of *D. variabilis* hemelipoglycoprotein 2 subunit with de novo sequenced protein from *Dermacentor marginatus*.

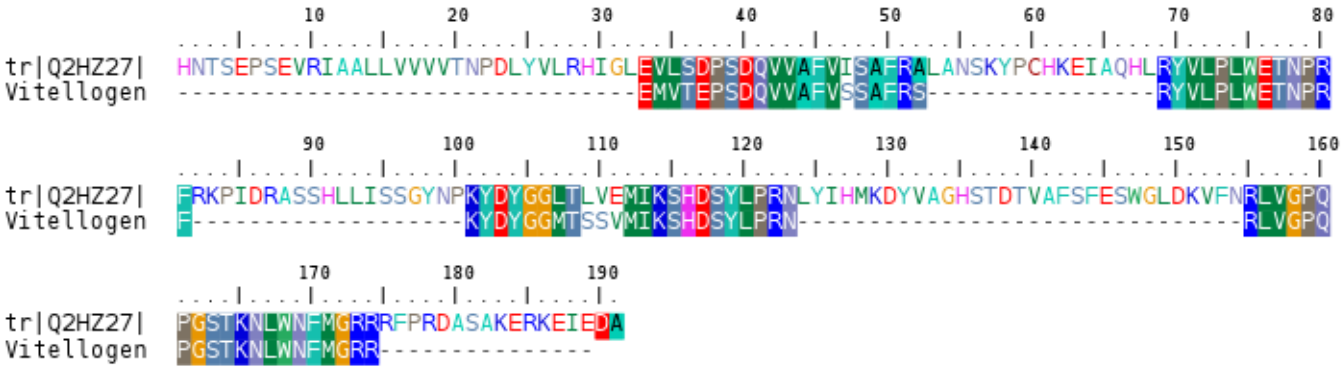

Figure 3.3: Comparison of *D. variabilis* vitellogenin fragment with de novo sequenated protein from *D. marginatus*.

Sequence fragment	HLGP 1		HLGP2		vitellogeinin	
	Match	Position	Match	Position	Match	Position
RMAALWALKQ	100,00%	[545-554]	100,00%	[489-498]		
KNLWNFMGRR	100,00%	[736-745]	100,00%	[680-689]	100,00%	[165-174]
RLVGPQPGSTKN	100,00%	[726-737]	100,00%	[671-681]	100,00%	[155-166]
KGVLSLFQLDLVKG	100,00%	[126-139]	100,00%	[71-84]		
RVAGNLATRK			90,00%	[460-469]		
REMVTEPSDQVVAFVSSAFRS	95,00%	[604-623]	100,00%	[547-567]	95,00%	[33-52]
RYVLPLWETNPRF	100,00%	[640-652]	76,00%	[584-596]	100,00%	[69-81]
KYDYGGMTSSVMIRS	80,00%	[672-686]	100,00%	[616-630]	80,00%	[101-155]
KSHDSYLPRN	100,00%	[685-694]	100,00%	[629-638]	100,00%	[114-123]
EVTKVKNLEK	90,00%	[175-184]	100,00%	[120-129]		
ESGYLFKPHGEG	83,00%	[237-248]	100,00%	[182-193]		
EGLKKLAHIEY	100,00%	[321-331]	81,00%	[265-275]		
HNLGALEQPFAAEVYQ			100,00%	[38-47]		

Table 3.2: Table of BLASTed fragments and their results, numbers in the square brackets denote the position of the fragment

The coverage per protein is depicted in the table 3.3.

	HLGP 1	HLGP2	vitellogenin
Coverage [%]	8.15	10.66	40.21

Table 3.3: Table of coverage per protein

The comparison of both hemelipoglycoprotein subunits and vitellogenin was done with ClustalW algorithm and can be seen in A.17 and A.18. The vitellogenin shows 100% matching identidy with HLGP 1 subunit in the middle (position 572-762). The similarity between vitellogenin and the HLGP

2 subunit is 71%. HLGP 2 subunit is shorter than HLGP 1 subunit by the first 55 amino acids. The match between both subunit is 67%.

3.1 Fragments and their sequences

The fragments were subjected to a BLAST database search with UniProt protein database for genus *Rhipicephalinae*. Those fragments that corresponded to hemelipoglycoprotein (see table 3.2) were aligned to complete sequences of both subunits of hemelipoglycoprotein and vitellogenin from *D. variabilis*. For the alignment, first step (pairwise alignment) of ClustalW algorithm was used to fit each fragment to corresponding place on complete reference sequence from *Dermacentor variabilis*. The results from BLAST database search can be seen in table 3.2 and the pairwise alignment results in figures 3.1, 3.2 and 3.3.

The m/z values for the fragments with the corresponding errors determined from the spectrum is not included in the Appendix for spectra A.2, A.4 and A.7 because of a harddisk hardware failure and inability to retrieve the pictures.

RMAALWALKQ

This sequence fragment showed 100% match for both subunits of hemelipoglycoprotein (HLGP). There are no differences in sequence between sequenced hemelipoglycoprotein from *D. marginatus* and reference sequence from *D. variabilis*. With no similarity to vitelogenin, this motif is unique for hemelipoglyco protein.

KNLWNFMGRR

This sequence fragment showed 100% match for both HLGP subunits and for vitellogenin. The sequence contains no changes and seems to be common for both subunits and vitellogenin.

RLVGPQGSTKN

The sequence shows again 100% match for both HLGP subunits and for vitellogenin. However, the sequence differs slightly in HLGP 2 subunit, where the first three amino acids of the fragment are RLV compared to HLL in corresponding position on the reference protein sequence. However, HLGP 1 subunit and vitellogenin do contain RLV motif that corresponds to the fragment and therefore I assume it as a part of HLGP first subunit peptide.

KGVLSLFQLDLVKG

This fragment shows 100% match for both HLGP subunits. There is no change in sequence between *D. marginatus* and *D. variabilis*.

RVAGNLATRK

This peptide shows only 90% match for HLGP 2 subunit. There is mutation from glycine (G) to alanine (A) on the position 466 (see fig: A.2). As there is no similarity to HLGP 1 subunit sequence, I conclude that this peptide is part of the HLGP 2 subunit with one amino acid difference between the two *Dermacentor* species.

REMVTEPSDQVVAFVSSAFRS

This sequence shows 100% match for HLGP 2 subunit and 90% for both vitellogenin and HLGP 1 subunit. Both vitellogenin and HLGP 1 subunit show difference in isoleucine (I) to serine (S) (as seen in figures 3.1 and 3.3). This means that the sequenced fragment is part of HLGP 2 subunit (see fig: A.5).

RYVLPLWETNPRF

The fragment indicates 100% match for first HLGP 1 subunit and vitellogenin and only 76% for HLGP 2 subunit. This shows that the sequence fragment corresponds to HLGP 1 subunit and that it is common motif between hemelipoglycoprotein and vitellogenin.

KYDYGGMTSSVMIRS

The sequence shows 80% match for HLGP 1 and vitellogenin and 100% match for HLGP 2. Therefore the fragment corresponds to HLGP 2 subunit. The spectrum is in A.6.

KSHDSYLPRN

The fragment indicates 100% match for both HLGP subunits and vitellogenin. There is no change in the sequence and it seems that the motif is common between hemelipoglyco protein and vitellogenin.

EVTKVKNLEK

The sequence shows 90% for HLGP 1 subunit and 100% match for HLGP 2 subunit. This indicates that the fragment corresponds to the HLGP 2 subunit.

ESGYLFKPHGEG

This fragment shows 83% match for HLPG 1 subunit and 100% match for HLGP 2 subunit. This clearly indicates that the sequence fragment corresponds to the second subunit HLGP 2.

EGLKKLAHIEY

This sequence matches 100% the HLGP 1 subunit. The result for HLGP 2 subunit is only 81%.Again this indicates that the sequence fragment corresponds to the HLGP 1 subunit.

GALEQPFAAE

This sequence is part of bigger fragment HLNGALEQPFAAEVYQ. The part GALEQPFAAE shows 100% match with HLGP 2 subunit with substitution lysine (K) to glutamic acid (E).

Chapter 4

Discussion

The obtained sequence fragments showed similarity to hemelipoglycoprotein subunits HLGP 1 and HLGP 2, as expected. Some of the fragments also showed similarity to vitellogenin. Since hemelipoglyco protein and vitellogenin share similarities, this observation is not surprising. Hemelipoglyco protein and vitellogenin are closely related on the primary structure level, but differ in posttranstaltional processing (Donohue et al., 2009).

4.1 Trypsin and V8 effectivity

One part of the sample was digested by *Staphylococus aureus* V8 enzyme (Glu-C) and second part by trypsin. Trypsin, however showed better results in terms of amount of successfully matched sequence fragments in comparison to V8 (see fig: 3.1). This may be caused by the use of bicarbonate buffer instead of phosphate buffer during digestion. In this buffer, V8 specificity changes from glutamic acid and aspartic acid to glutamic acid only (Drapeau et al., 1972). In comparison, trypsin sensitivity to lysine and arginine (unless followed by proline) is not affected by the presence of phosphate ions. Hence, according to graphs in section A.2, trypsin would be more advantageous choice than V8 in a phosphate lacking enviroment in case of digestion of hemelipoglycoprotein.

4.2 Coverage

The sequencing of HLGP resulted in coverage of 8.15% for HLGP 1 subunit and 10.66% for subunit HLGP 2. Since HLGP is closely related to vitellogenin on the primary structure level, the results also include coverage for vitellogenin. The known fragment of vitellogenin in the database is fairly small and therefore the coverage is relatively high (see table 3.3).

The theoretical maximum coverage for HLGP 1 subunit is 87.34% with trypsin, 95.75% with V8 in bicarbonate buffer and 82.57% with V8 in phosphate buffer. The theoretical coverage for HLGP 2 subunit is 89.38% with trypsin, 86.71% with V8 in bicarbonate buffer and 84.19% with V8 in phosphate buffer. By combining trypsin and V8 enzymes, it is theoretically possible to yield 99.42%

coverage for HLGP 1 and 99.19% for HLGP 2. By taking into account the missed-cleavage sites, it would be theoretically possible to obtain the whole sequence of the protein. The enzyme does not work with 100% effectivity, especially when there are lots of cleavage sites close together.

Interestingly, in case of the HLGP 1 subunit, theoretically V8 in phosphate buffer would be the best choice. On the other hand, the theoretical coverage in case of HLGP 2 subunit supports the real results, which favored trypsin to V8.

The real coverage is approximately 10 times lower that the theoretical values. This may be caused by precipitation of the protein in the sample or incomplete digestion. To elevate the coverage percentage, it would be ideal to perform the V8 digest in phosphate buffer too (since the thetheoretical coverage for HLGP 1 subunit is highest with phosphate buffer). Another possibility would be optimalization of the digestion reaction by finding the best ratio protein/enzyme for the reaction. According to Sigma Product Information bulletin, the optimal ratio is between 1:100 and 1:20 of enzyme to substrate.

During performance of 1 analysis with optimal conditions, the resulting coverage is rarely over 50%. Therefore to elevate the success of the sequenation, it is necessary to perform the analysis multiple times.

4.3 Sequence fragments

Almost all the sequence fragments showed 100% match at least to one of HLGP subunits. Some of them showed similarities to vitellogenin. This is not surprising, because vitellogenin and HLGP subunit show 100% match in the middle of the subunit and the HLGP subunits themselves show 67% match between them.

However, there are 2 fragments, where the results need to be explained. First, there is the sequence RVAGNLATRK, that shows 90% match only to HLGP 2 subunit (see table 3.2 and figure A.2. The results show alanine (A) on the position 466, whereas the HLGP 2 subunit sequence in the database indicates presence of glycine (G). The difference may be caused by variability between individual ticks,variability between the species or error in the database. The change from G to A will not cause change in the cleavage or the presence of an active group and also the steric consequences will be minimal. Thus, the conformation and function of the protein will not be changed. This might indicate first found difference between HLGP from *D. marginatus* and *D. variabilis*.

Second fragment that needs to be explained is HNLGALEQPFAAEVYQ. There is only part of the fragment that shows 100% match with HLGP 2 subunit (see table 3.2). The matching sequence is GAELQPFAAE, with change from lysine (K) to glutamic acid (E). The quality of the spectra however can not fully prove such difference. The glutamic acid is in both y and b fragmentation series as a part of a not fully fragmented cluster of amino acids. Furthermore the effect of such change from basic to acidic amino acid can not be fully clarified without additional information.

Chapter 5

Conclusion

The aim of the work was characterization of hemelipoglycoprotein from tick *Dermacentor marginatus* plasma by mass spectrometry.

The protein was sequenced by *de novo* sequenation approach by tandem MS and yielded sufficient ammount of sequence fragments that allowed identification and subsequent partial sequence characterization of the protein. Both subunits of the tick heme storage protein were sequenced with similar coverage of 8.15% for first subunit and 10.66% for second subunit. The sequenced fragments also represent 40.21% of a known tick vitellogenin fragment, which was not surprising, because the hemelipoglycoprotein shares similarities with vitellogenin in terms of primary structure (Donohue et al., 2009).

The experiments showed, that better results were obtained with trypsin, than with V8 as an enzyme of choice for protein digestion before MS/MS analysis.

Two putative differences between HLGP sequences of *D. marginatus* and *D. variabilis* were discovered. Both of them are in the HLGP 2 subunit. One difference is found in the position 47, but has to be confirmed. The second change can be found in the position 466 and seems plausible.

The aim of further study of the protein by mass spectrometry would be to obtain more information about the primary structure and the modifications of the protein, especially N- and O-glycosylations.

Appendix A

Appendix - spectra and sequences

A.1 Spectra of peptide fragments

a	136,08	235,14	348,23	445,28	558,37	744,44	873,49	974,54	1088,58	1185,63	--
	(0)	(0)		(-0,01)	(-0,01)	(0)					
b	164,07	263,14	376,22	473,28	586,36	772,44	901,48	1002,53	1116,57	1213,63	--
		(0)	(0)	(0)	(-0,01)	(-0,01)	(-0)				
	Y	V	L	P	L	W	E	T	N	P	R
y"	--	1224,67	1125,61	1012,52	915,47	802,38	616,31	487,26	386,22	272,17	175,12
		(-0)	(0)	(0)	(-0)	(-0)	(0)	(-0)	(0)	(0)	(0)
z	--	1207,65	1108,58	995,49	898,44	785,36	599,28	470,24	369,19	255,15	158,09
			(-0,03)	(-0,02)		(-0)	(-0,02)	(-0)	(-0)	(0)	(0)

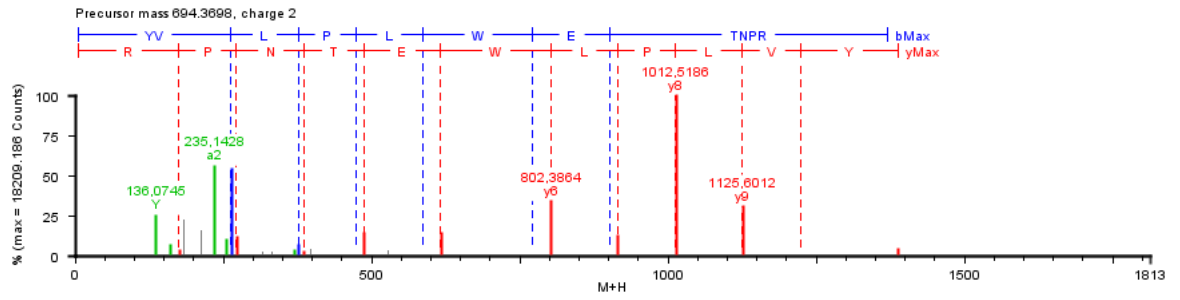

Figure A.1: MS/MS spectrum peptide RYVLPLWETNPRF and m/z values for the fragments with the corresponding errors as determined from spectrum.

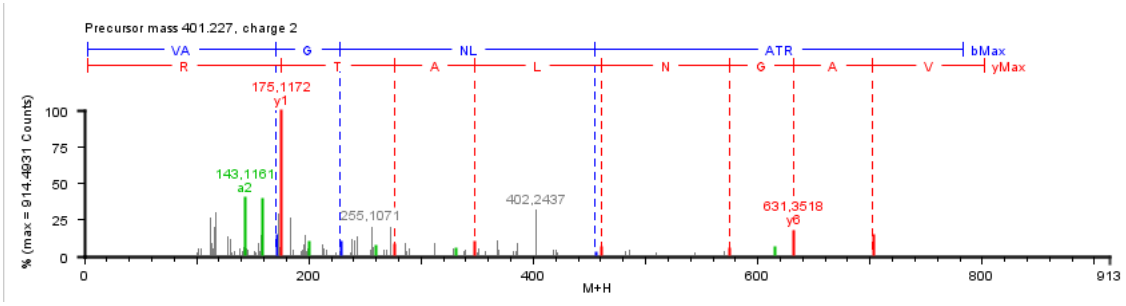

Figure A.2: MS/MS spectrum peptide RVAGNLATRK.

a	120,05	191,09	262,12	375,21	561,29	632,32	745,41	--
	(-0,03)	(0)	(0)					
b	148,04	219,08	290,12	403,2	589,28	660,32	773,4	--
		(0)	(0)	(0,01)			(-0,07)	
	M	A	A	L	W	A	L	K
y''	--	772,47	701,43	630,4	517,31	331,23	260,2	147,11
		(0)	(-0,01)	(0)	(-0)	(0)	(0)	(0)
z	--	755,45	684,41	613,37	500,29	314,21	243,17	130,09
							(0)	(0)

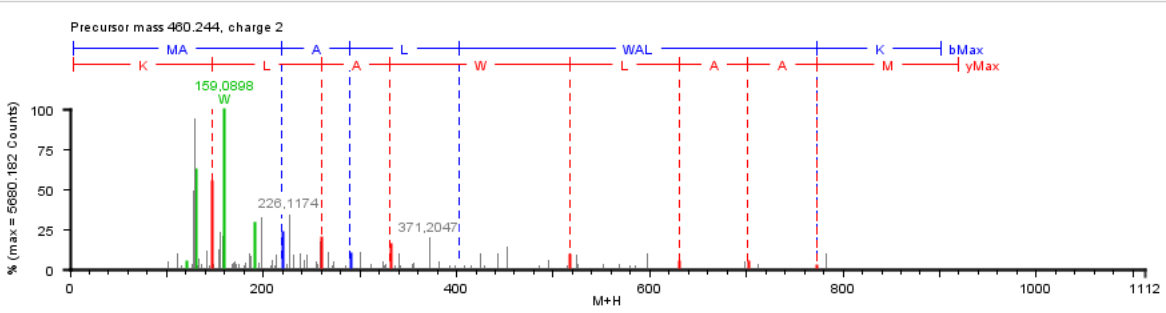

Figure A.3: MS/MS spectrum peptide RMAALWALKQ and m/z values for the fragments with the corresponding errors as determined from spectrum.

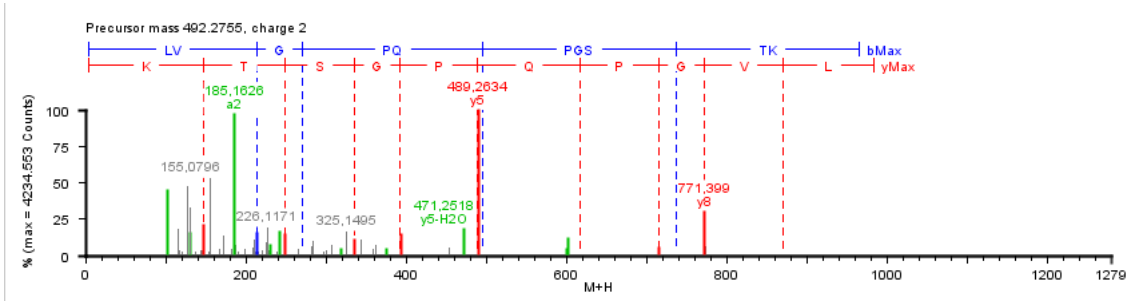

Figure A.4: MS/MS spectrum peptide RLVGPQPGSTKN.

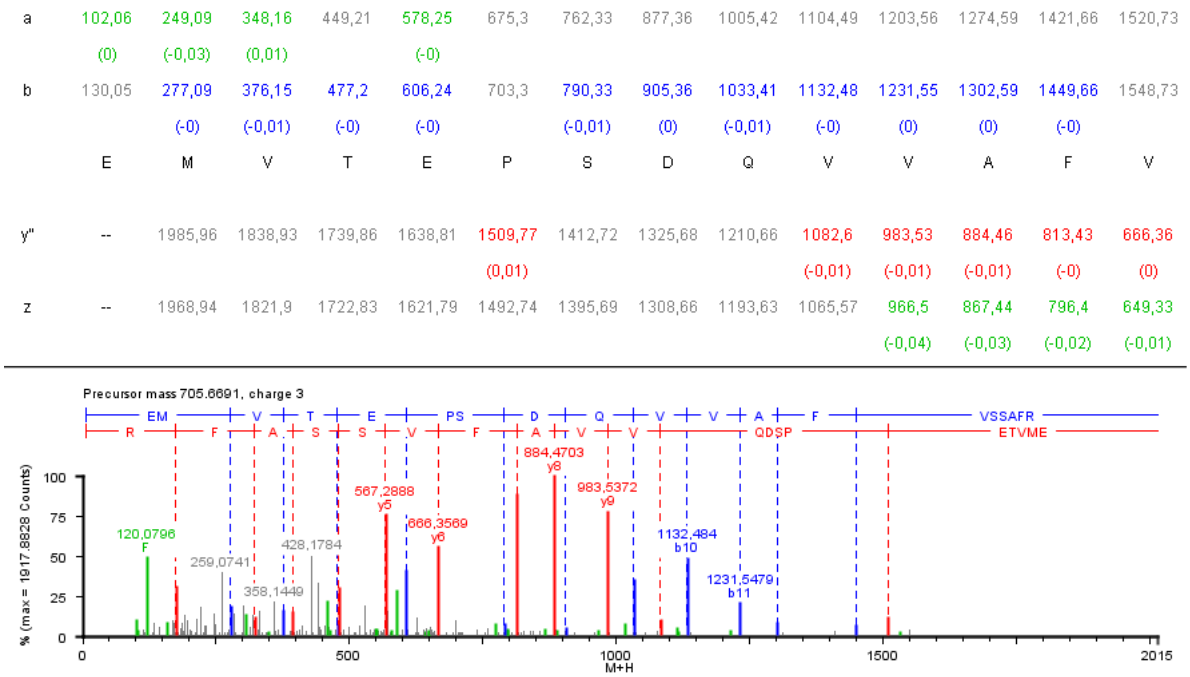

Figure A.5: MS/MS spectrum peptide REMVTEPSDQVVAFVSSAFRS and m/z values for the fragments with the corresponding errors as determined from spectrum.

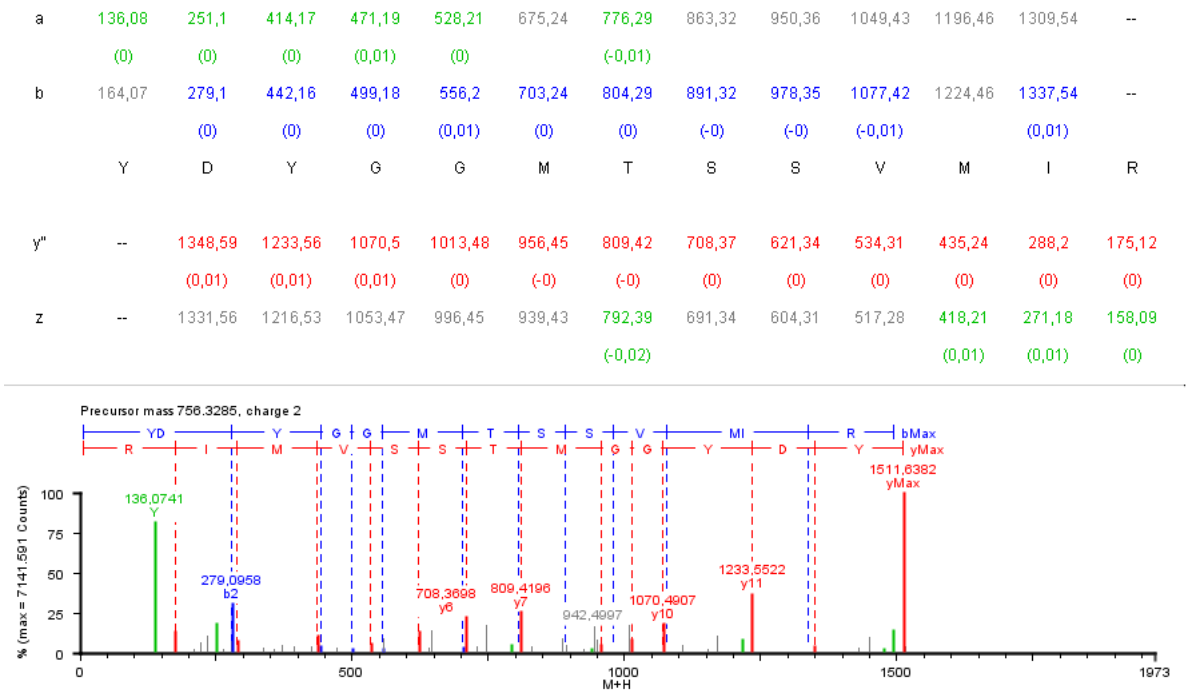

Figure A.6: MS/MS spectrum peptide KYDYGGMTSSVMIRS and m/z values for the fragments with the corresponding errors as determined from spectrum.

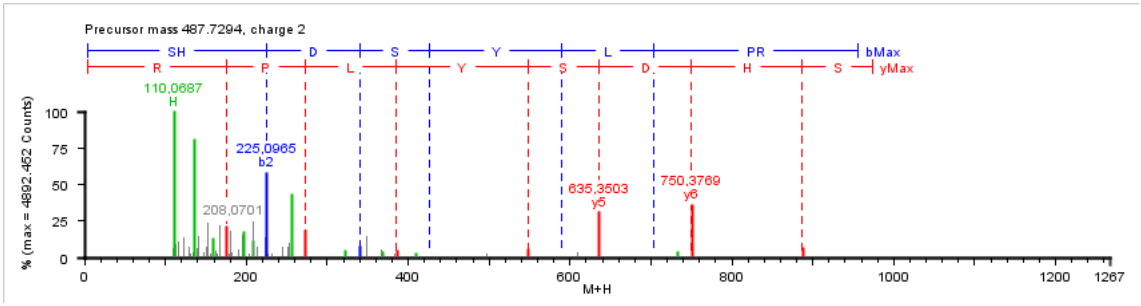

Figure A.7: MS/MS spectrum peptide KSHDSYLPRN.

a	87,06	200,14	386,22	500,26	647,33	778,37	835,39	--
		(-0)						
b	115,05	228,13	414,21	528,26	675,33	806,37	863,39	--
		(0,01)						
	N	L	W	N	F	M	G	R
y''	--	923,46	810,37	624,29	510,25	363,18	232,14	175,12
			(-0,02)	(-0)	(-0)	(-0,01)	(0)	(0)
z	--	906,43	793,35	607,27	493,22	346,15	215,11	158,09
						(-0,01)	(0)	

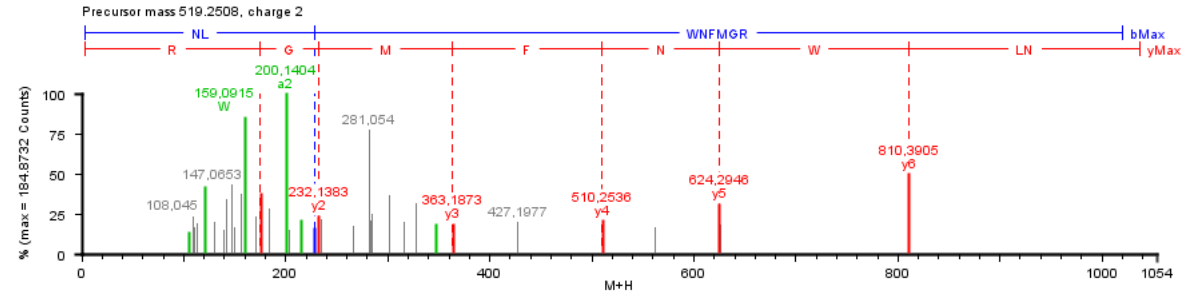

Figure A.8: MS/MS spectrum peptide KNLWNFMGRR and m/z values for the fragments with the corresponding errors as determined from spectrum.

a	30,03	129,1	242,19	329,22	442,3	589,37	717,43	830,51	945,54	1058,62	1157,69	--
		(0)	(0,03)	(0)	(0,01)	(0,03)	(0,01)					
b	58,03	157,1	270,18	357,21	470,3	617,37	745,42	858,51	973,54	1086,62	1185,69	--
		(0)	(0)	(0,02)	(0,03)	(0,03)						
	G	V	L	S	L	F	Q	L	D	L	V	K
y''	--	1274,77	1175,7	1062,62	975,59	862,5	715,44	587,38	474,29	359,27	246,18	147,11
			(-0,01)	(-0)	(-0)	(0)	(-0)	(0)	(0)	(0)	(0)	(0)
z	--	1257,75	1158,68	1045,59	958,56	845,48	698,41	570,35	457,27	342,24	229,16	130,09
							(0,01)		(0,01)	(0,04)	(0,04)	(0)

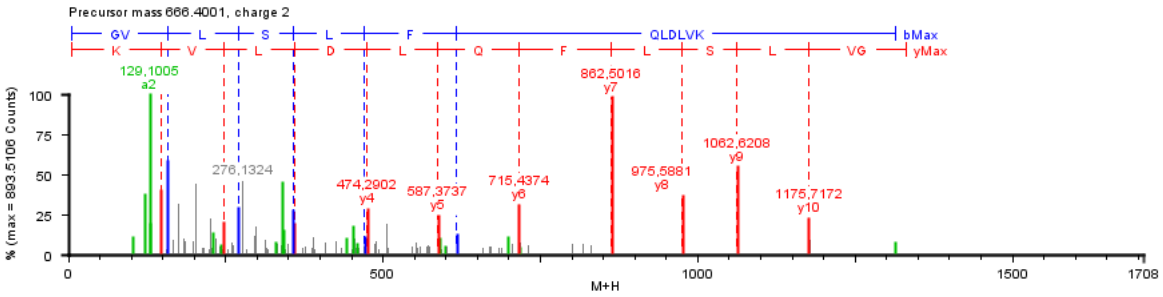

Figure A.9: MS/MS spectrum peptide KGVLSLFQLDLVKG and m/z values for the fragments with the corresponding errors as determined from spectrum.

a	72,08	173,13	301,22	400,29	528,39	642,43	755,51	--
		(-0)						
b	100,08	201,12	329,22	428,29	556,38	670,43	783,51	--
		(0)	(-0)	(-0,01)	(-0,02)			
	V	T	K	V	K	N	L	E
y''	--	831,49	730,45	602,35	503,28	375,19	261,15	148,06
		(0,02)	(-0,01)	(-0,01)	(-0)	(-0,01)	(-0)	(0)
z	--	814,47	713,42	585,32	486,26	358,16	244,12	131,03
			(-0,02)	(-0,02)	(-0,01)	(0)	(-0)	

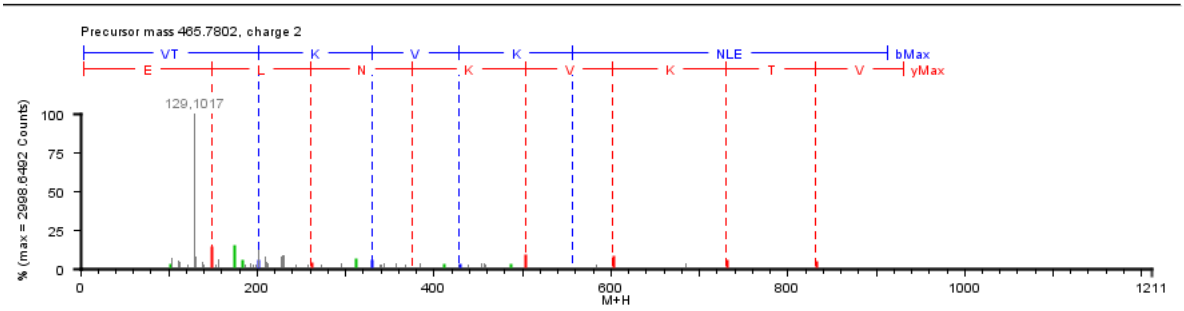

Figure A.10: MS/MS spectrum peptide EVTKVKNLEK and m/z values for the fragments with the corresponding errors as determined from spectrum.

a	60,04	117,07	280,13	393,21	540,28	668,38	765,43	902,49	959,51	--
		(0)	(0)	(-0,02)			(0,03)	(0,02)	(0,02)	
b	88,04	145,06	308,12	421,21	568,28	696,37	793,42	930,48	987,51	--
		(0)	(0)	(0,03)	(-0,01)	(0,02)		(0,01)		
	S	G	Y	L	F	K	P	H	G	E
y"	--	1047,53	990,5	827,44	714,36	567,29	439,19	342,14	205,08	148,06
				(-0)	(0)	(-0)	(0)	(0)	(0)	(-0)
z	--	1030,5	973,48	810,41	697,33	550,26	422,17	325,11	188,06	131,03
						(-0,02)		(-0,1)		(-0,02)

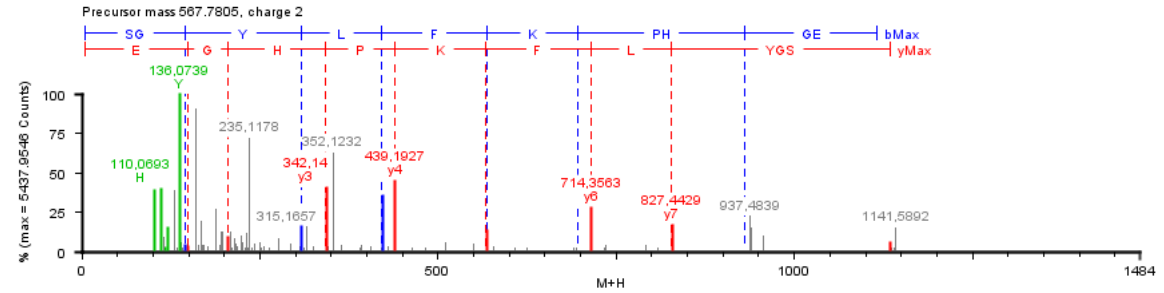

Figure A.11: MS/MS spectrum peptide ESGYLFKPHGEG and m/z values for the fragments with the corresponding errors as determined from spectrum.

a	30,03	143,12	271,21	399,31	512,39	583,43	720,49	833,57	--
		(0)					(0)	(0)	
b	58,03	171,11	299,21	427,3	540,39	611,42	748,48	861,57	--
		(-0)					(0)	(0,01)	
	G	L	K	K	L	A	H	I	E
y"	--	951,6	838,51	710,42	582,33	469,24	398,2	261,14	148,06
			(-0,02)	(0)	(-0,01)	(-0,01)	(-0)	(0)	(-0)
z	--	934,57	821,49	693,39	565,3	452,21	381,18	244,12	131,03

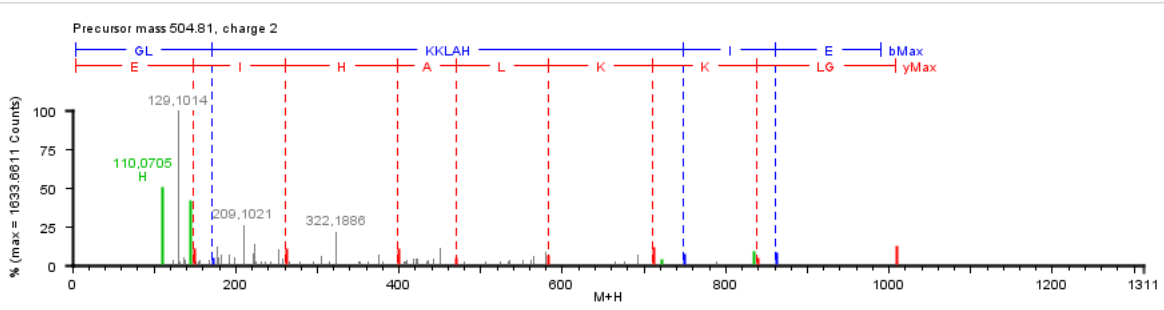

Figure A.12: MS/MS spectrum peptide EGLKKLAHIEY and m/z values for the fragments with the corresponding errors as determined from spectrum.

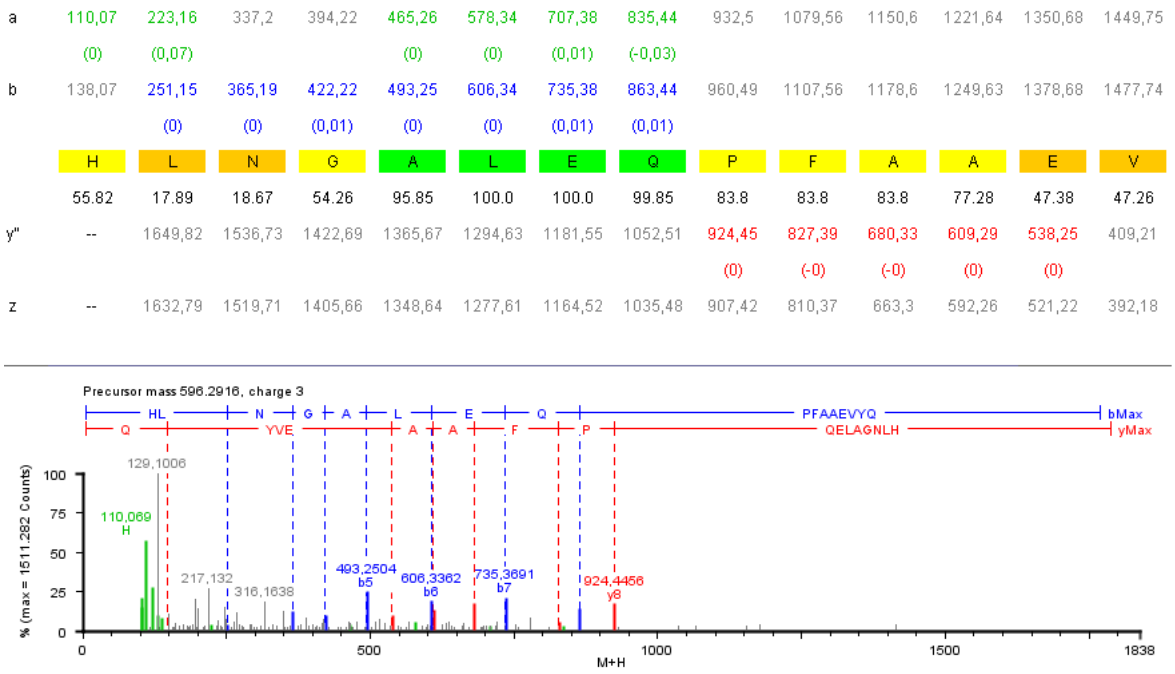

Figure A.13: MS/MS spectrum peptide HLNGALEQPFAAEVYQ and m/z values for the fragments with the corresponding errors as determined from spectrum.

A.2 Amino acid composition of hemelipoglycoprotein

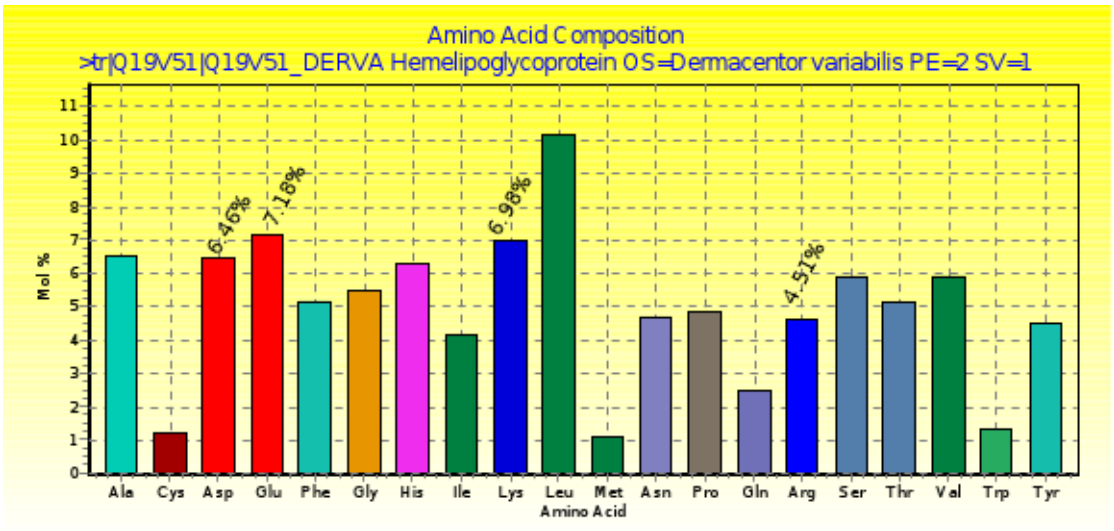

Figure A.14: Amino acid composition of HLGP first subunit (HLGP 1)

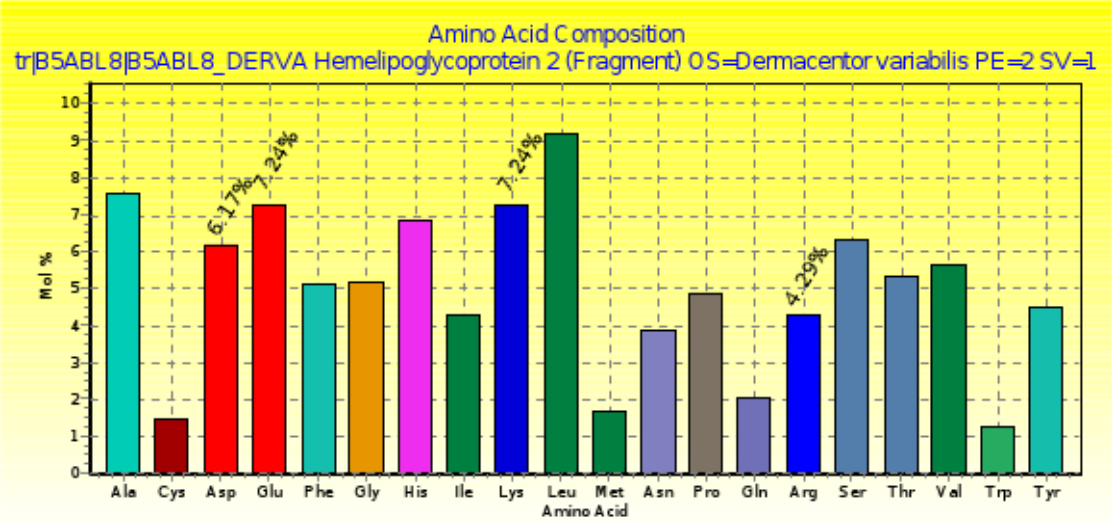

Figure A.15: Amino acid composition of HLGP second subunit (HLGP2)

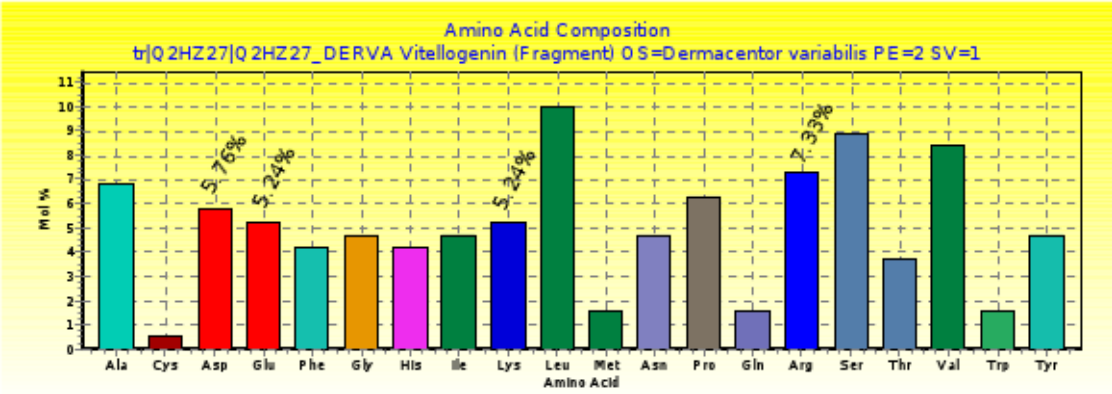

Figure A.16: Amino acid composition of vitellogenin fragment

A.3 Hemelipoglycoprotein and vitellogenin sequences

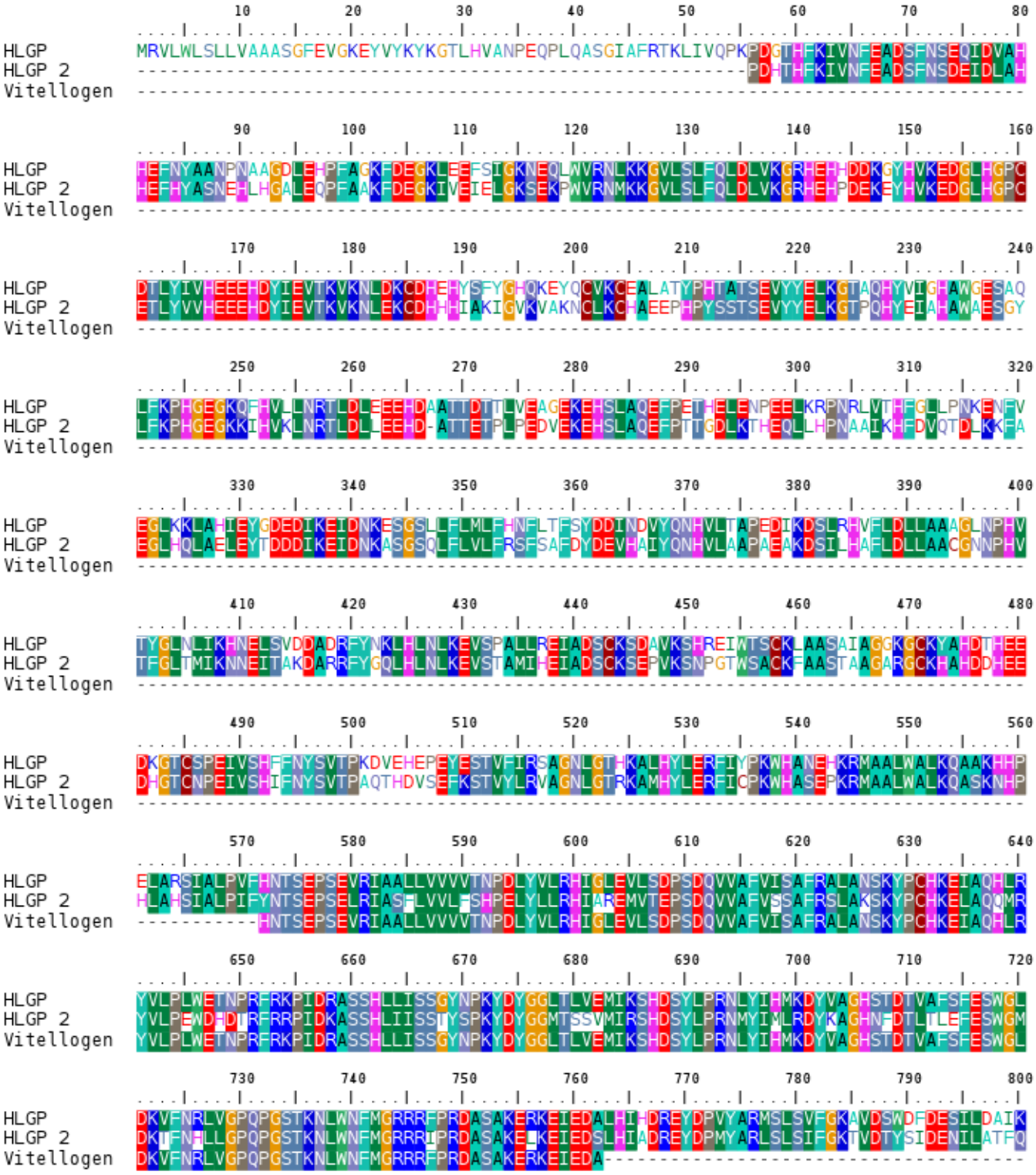

Figure A.17: Alignment of HLGP subunits and Vitellogenin fragment, ClustalW algorithm, part 1.

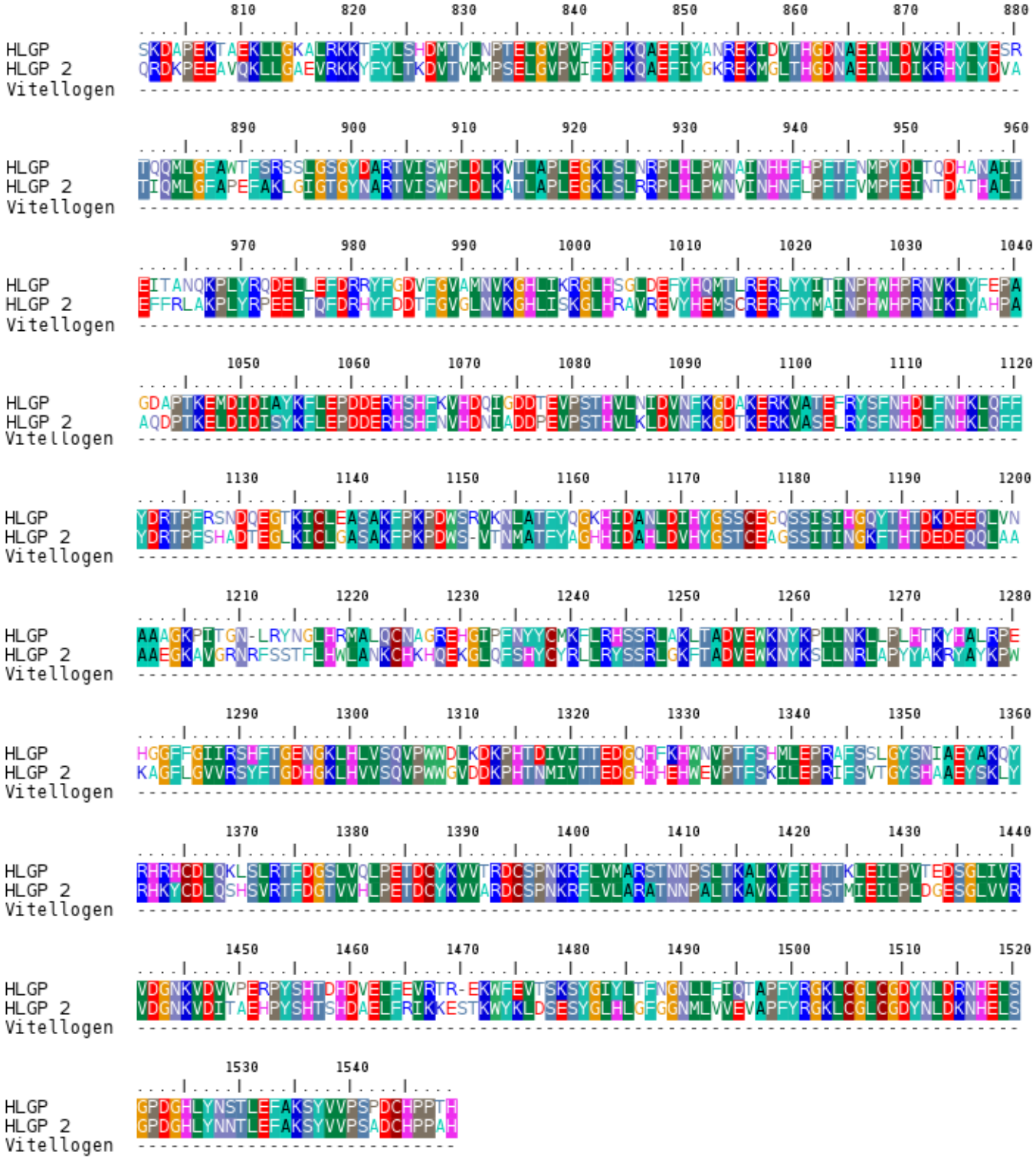

Figure A.18: Alignment of HLGP subunits and Vitellogenin fragment, ClustalW algorithm, part 2.

Figure A.19: Digestion map of HLGP subunit, trypsin and V8 enzymes, black dot denotes trypsin, red square denotes V8.

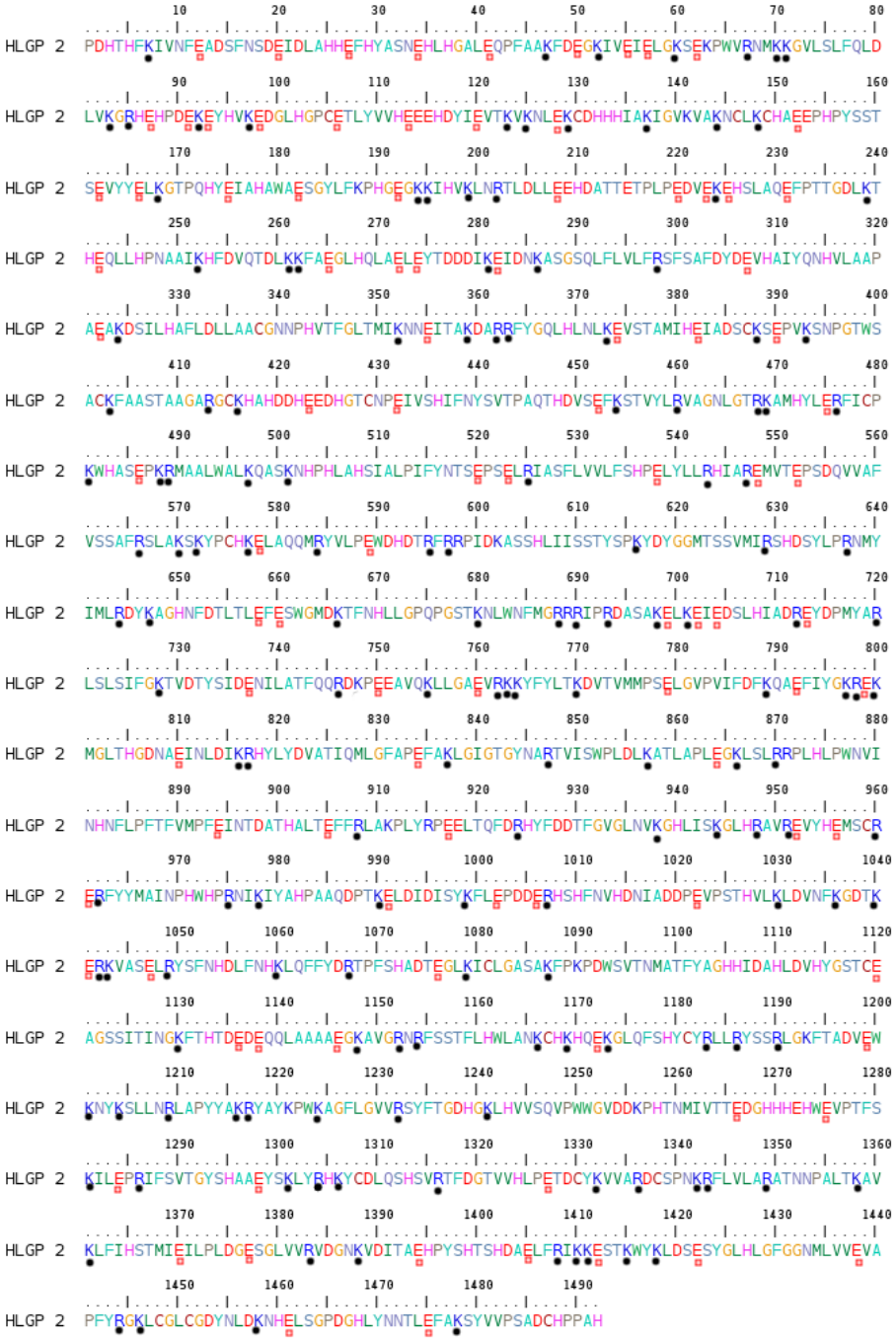

Figure A.20: Digestion map of HLGP 2 subunit, trypsin and V8 enzymes, black dot denotes trypsin, red square denotes V8.

Bibliography

Beati, L. and R. Raoult (1993). *Rickettsia massiliae* sp. nov, a new spotted fever group of rickettsia. *Int. J. Syst. Bacteriol. 43*, 839–340.

Donohue, K., S. Khalil, R. Mitchell, D. Sonenshine, and R. M. Roe (2008). Molecular characterization of the major hemelipoglycoprotein in ixodid ticks. *Insect Mol. Biol. 17*, 197–208.

Donohue, K., S. Khalil, D. Sonenshine, and R. M. Roe (2009). Heme-binding storage proteins in the chelicerata. *J Insect Physiol. 55*, 287–296.

Drapeau, G. R., Y. Boily, and J. Houmard (1972). Purification and properties of and extracellular protease of *Staphylococus aureus*. *J. Biol. Chem. 20*, 6720–6726.

Dupejová, J. (2008a). Izolace a biochemická charakteristika proteinů frep u klíštěte *Dermacentor marginatus*. Bachelor thesis, Faculty of Science, University of South Bohemia, České Budějovice, p. 8.

Dupejová, J. (2008b). Izolace a biochemická charakteristika proteinů frep u klíštěte *Dermacentor marginatus*. Bachelor thesis, Faculty of Science, University of South Bohemia, České Budějovice, p. 29-31.

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

Hall, T. A. (1999). Bioedit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/nt. *Nucl. Acids. Symp. Ser. 41*, 95–98.

Hillyard, P. D. (1996). *Ticks of North-West Europe*. FSC Publications, ISBN 1 85153 257 9, pages: 3; 22; 116.

Johnson, R. S., S. A. Martin, K. Biemann, J. T. Stults, and J. T. Watson (1987). Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: Differentiation of leucine and isoleucine. *Anal. Chem. 59*, 2621–2625.

Kovář, L. (2004). Tick saliva in anti-tick immunity and pathogen transmission. *Folia Microbiol. 49*, 327–336.

Mamyrin, B. A., V. I. Karataev, D. V. Shmikk, and V. A. Zagulin (1973). The mass-reflectron, a new non-magnetic time-of-flight mass spectrometer with high resolution. *Sov. Phys. JETP 37*, 45–48.

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

Maya-Monteiro, C. M., S. Daffre, C. Logullo, F. A. Larra, E. W. Ales, and M. L. Capurro (2000). Help, a heme lipoprotein from the hemolymph of the cattle tick, *Boophilus microplus*. *J. Biol. Chem. 275*, 36584–36589.

Noga, M., T. Dylag, and J. Silberring (2009). Sequencing of peptides and proteins. In R. Eckman, J. Silberring, A. Westman-Brinkmalm, and A. Kraj (Eds.), *Mass Spectrometry: in instrumentation, interperetation and applications*, pp. 178–210. Jonh Wiley & Sons, Inc., Hoboken, New Jersey.

Nuttall, P. A. and M. Labuda (2004). Tick-host interactions: saliva-activated transmission. *Parasitology 129 Suppl.*, S177–189.

Nuttall, P. A. and M. Labuda (2008). Saliva-assisted transmission of tick-borne pathogens. In A. S. Bowman and P. Nuttall (Eds.), *Ticks: Biology, Disease and Control*, pp. 205–219. Cambridge University Press.

Paul, W. and H. Steinwedel (1953). Ein neues massenspektrometer ohne magnetfeld. *Z. Naturforch. A8*, 448–450.

Roepstorff, P. and J. Fohlman (1984). Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom. 11*, 601.

Sonenshine, D. E. (1991). *Biology of ticks Vol. 1.* Oxford University Press, New York.

Sonenshine, D. E. (1993). *Biology of ticks Vol. 2.* Oxford University Press, New York.

Strohalm, M., D. Kavan, P. Novák, M. Volný, and V. Havlíček (2010). mmass 3: A cross-platform software environment for precise analysis of mass spectrometric data. *Anal. Chem. (In Press)*.

Westman-Brinkmalm, A. and G. Brinkmalm (2009a). A mass spectrometer's building blocks. In R. Eckman, J. Silberring, A. Westman-Brinkmalm, and A. Kraj (Eds.), *Mass Spectrometry: in instrumentation, interperetation and applications*, pp. 2–87. Jonh Wiley & Sons, Inc., Hoboken, New Jersey.

Westman-Brinkmalm, A. and G. Brinkmalm (2009b). Tandem mass spectrometry. In R. Eckman, J. Silberring, A. Westman-Brinkmalm, and A. Kraj (Eds.), *Mass Spectrometry: in instrumentation, interperetation and applications*, pp. 88–103. Jonh Wiley & Sons, Inc., Hoboken, New Jersey.

Wikipedia (2010a). Microchannel plate detector. `http://upload.wikimedia.org/wikipedia/commons/5/50/Mcp_schematic.gif`. only picture reprinted.

Wikipedia (2010b). Quadrupole mass analyzer. http://upload.wikimedia.org/wikipedia/commons/a/a7/Quadrupole_en.gif. only picture reprinted.

Wikipedia (2010c). Tandem mass spectrometry. http://upload.wikimedia.org/wikipedia/commons/f/fb/Peptide_fragmentation.gif. only picture reprinted.

Řeháček, J., J. Urvölgyi, and E. Kováčová (1977). Massive occurrence of rickettsiae of the spotted fever group in fowl tampan, argas persicus, in the armenian s.s.r. *Acta Virol. 21*, 431–438.