Two-hybrid analysis and functional annotation of Bm86 and ATAQ from *Rhipicephalus microplus*

By

Ilkadim Kiper

Supervisor: Dr. Christine Maritz-Olivier

Submitted in partial fulfilment of the requirements for the degree

MAGISTER SCIENTIAE

In the Faculty of Biological & Agricultural Sciences
School of Biological Sciences
Department of Genetics
University of Pretoria
South Africa

November 2013

Two-hybrid analysis and functional annotation of Bm86 and ATAQ from *Rhipicephalus microplus*

By

Ilkadim Kiper

Supervisor: Dr. Christine Maritz-Olivier
Anti-tick vaccine Development
Department of Genetics
University of Pretoria
Pretoria

Department of Genetics

Magister Scientiae

Dissertation abstract

Vaccination with recombinant Bm86 was found to protect against tick infestation, however, the efficacy of the vaccine varies against different tick species and geographical strains. An improvement of current vaccines could be achieved by identifying homologs of the current Bm86 antigen. A novel protein ATAQ was previously identified during a study that aimed to identify Bm86 homologs. It was shown that these two proteins do not share identical antigenic and immunogenic regions and that they also have different expression patterns. This indicates that ATAQ is not a Bm86 homolog, but rather an entirely different protein. Hence it is reasonable to assume that Bm86 antibodies would not cross-react with ATAQ antigens. Determination of the function of Bm86 and its structurally related protein ATAQ would provide invaluable insight into an unexplored biological system in ticks and potentially lead to the development of an improved vaccine formulation. Therefore, this study sets out to analyze protein-protein interactions by means of the yeast two-hybrid screening system using Bm86 and ATAQ from *R. microplus* as bait proteins.

We were able to obtain that Bm86 and ATAQ have two common (aldehyde and retinol dehydrogenases), as well as a unique (Kunitz-like protein) interacting partners, respectively, via yeast two-hybrid screening. Further confirmation was achieved by immunoprecipitation, western blot and LC/MS analysis.

The aldehyde and retinol dehydrogenases, as well as the Kunitz-like protein are involved in retinol metabolism. Therefore, as Bm86 and ATAQ share common binding partners, we hypothesized that they both are involved in the same metabolic pathway in the ticks. Future studies, would involve further confirmation of this pathway in *R. microplus* and evaluation of a mixture of both Bm86 and its interacting partners in animal vaccination trials.

Declaration

I, ... declare that this dissertation, which I hereby submit for the degree ..... at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature: …… Date: ……

Plagiarism declaration:

Full names of student: fill in

Student number: fill in

Declaration

1. I understand what plagiarism is and am aware of the University's policy in this regard.
2. I declare that this dissertation is my own original work. Where other people's work has been used (either from a printed properly acknowledged and referenced in accordance with the departmental requirements.
3. I have not used work previously produced by another student or any other person to hand in as my own.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature of student:

Date: __

Acknowledgements

Firstly I would like to thank my supervisor Dr. Christine Maritz-Olivier for her guidance and knowledge that have inspired me as a scientist.

I would like to thank my sister Ilkser E. Kiper and my friend Sabine A. Richards for their invaluable advice, support and encouragement.

I would like to thank the Red Meat Research Development Trust of SA, the Technology and Human Resources for Industry Programme (THRIP), and the National Research Foundation (NRF) for funding the project. I would like to also thank Dr A. Nijhof for providing ATAQ clones and Prof Jose De la Fuente for providing recombinant Bm86. Furthermore, I would like to thank ClinVet, Bloemfontein, South Africa and The University of Pretoria Biomedical Research Centre (UPBRC) for tick rearing. I would also like to thank Mariette Botha for the library construction and her advice.

I would like to thank all my friends for their help, assistance and friendship, especially Willie A. Van Zyl and Christian Stutzer.

I would especially like to thank my parents and family who have always encouraged me to follow my goals and for always being there for me whenever I needed them.

Finally, I would like to thank my fiancé Reynier Rautenbach, without him and his support I would have been lost during the difficult times of my studies.

Table of contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

A……..	Alanine
AD……	Activation Domain
Ade …….	Adenine
BAC……	Bacterial Artificial Chromosome
BD…….	Binding Domain
BLAST……	Basic Local Alignment Search Tool
bp…….	Base Pairs
BmiGI…….	A database of cDNAs expressed in *Boophilus microplus,* the ***B****oophilus* ***mi****croplus* **G**ene **I**ndex
ºC……	Degrees Celcius
C……..	Cystein
CAPS……..	3-(Cyclohexylamino)-1-Propanesulfonic Acid
CCB…….	Colloidal Coomassie Blue
cDNA……..	Complementary DNA
cfu……..	Colony Forming Units
C-terminal……	Carboxy Terminal
D……..	Aspartic Acid
Da……	Dalton
DAAF……..	Department of Agriculture, Fisheries and Forestry
DDO……	Double Dropout
DNA…….	Deoxyribonucleic Acid
DNA-BD……	DNA-Binding Domain
dNTP……...	Deoxynucleotide Triphosphate
DO……..	Dropout
DR……...	Disordered Regions

ds……	Double Stranded
E……..	Glutamic Acid
E. coli……..	*Escherichia coli*
EDTA……..	Ethylene Diamine Tetra Acetic Acid
EGF…….	Epidermal Growth-factor
ELISA…….	Enzyme-Linked Immunosorbent Assay
EST…….	Expressed Sequence Tag
F……..	Phenyalanine
FAO…….	The State of Food and Agriculture
FDR…….	False discovery rate
G…….	Glycine
GAL4 …….	Galactose 4 Regulatory Protein
Gbp…….	Giga Base Pairs
GBA……	Guilt-By-Association
GDP……	Gross Domestic Product
GPI……..	Glycosylphosphatidylinositol
I……	Isoleucine
IPTG……	Isopropyl-β-D-Thiogalactopyranoside
Ig…….	Immunoglobulin
K……..	Lysine
kDa…….	Kilo Dalton
L…….	Leucine
lacZ……..	β-Galactosidase gene
LC-MS…….	Liquid Chromatography–Mass Spectrometry
LB……	Luria-Bertani
M……..	Methionine
MCS……	Multiple Cloning Site
mRNA…….	Messenger RNA

N……..	Asparagine
NCBI……	National Centre for Biotechnology Information
N-terminal……..	Amino Terminal
PAGE……..	Polyacrylamide Gel Electrophoresis
PBS…….	Phosphate-Buffered Saline
PCR…….	Polymerase Chain Reaction
PDB…….	Protein Data Base
PHYRE……...	Protein Homology/analogy Recognition Engine
PPI……..	Protein-Protein Interactions
PSI-BLAST……	Position-Specific Iterative-BLAST
PVDF……..	Polyvinylidene Fluoride
Q…….	Glutamine
QDO……..	Quadruple Dropout
R……..	Arginine
RACE…….	Random Amplification of cDNA Ends
RNA…….	Ribonucleic Acid
RNAi…….	RNA interference
S…….	Serine
SCOP…….	**S**tructural **C**lassification **o**f **P**roteins database
SD……..	Standard Dropout
SDS…….	Sodium Dodecyl Sulfate
T…….	Threonine
TAE…….	Tris-acetate EDTA buffer
Taq…….	*Thermus aquaticus*
TBS…….	Tris Buffered Saline
TDO…….	Triple Dropout
TEMED…….	Tetramethylethylenediamine
Tm……..	Melting Temperature
Tris…….	Tris(hydroxymethyl) aminomethane
UAS…….	Upstream Activation Sequence

V…….. Valine

W…… Tryptophan

X…….. Any Amino Acid

X-gal…….. 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Y…….. Tyrosine

Chapter 1: Literature Review

1. 1. INTRODUCTION: TICKS AND TICK-BORNE DISEASES

1. 1. 1. Tick Taxonomy

Ticks are obligate hematophagous ectoparasites, capable of feeding on almost all types of terrestrial vertebrates including mammals, birds, reptiles and amphibians (Hovius *et al.*, 2008). They belong to the class Arachnida, subclass Acari, order Parasitiformes and the suborder Ixodida, which contains three families (Sauer *et al.*, 2000). The family Argasidae (soft ticks) contains 5 genera and more than 180 species, the Ixodidae (hard ticks) consists of 683 species and finally the Nuttalliellidae containing only one species called *Nuttalliella namaqua* (Barker, Murrell, 2004; Mans *et al.*, 2011; Mans *et al.*, 2004). The evolutionary relationship between the different tick families and selected subfamilies are indicated in Figure 1.1. It is suggested that ticks have evolved during the Cretaceous period, some 120 million years ago (MYA), or even as long ago as the Devonian period, some 400 MYA (Dobson, Barker, 1999; Klompen *et al.*, 1996; Kuno, Chang, 2005).

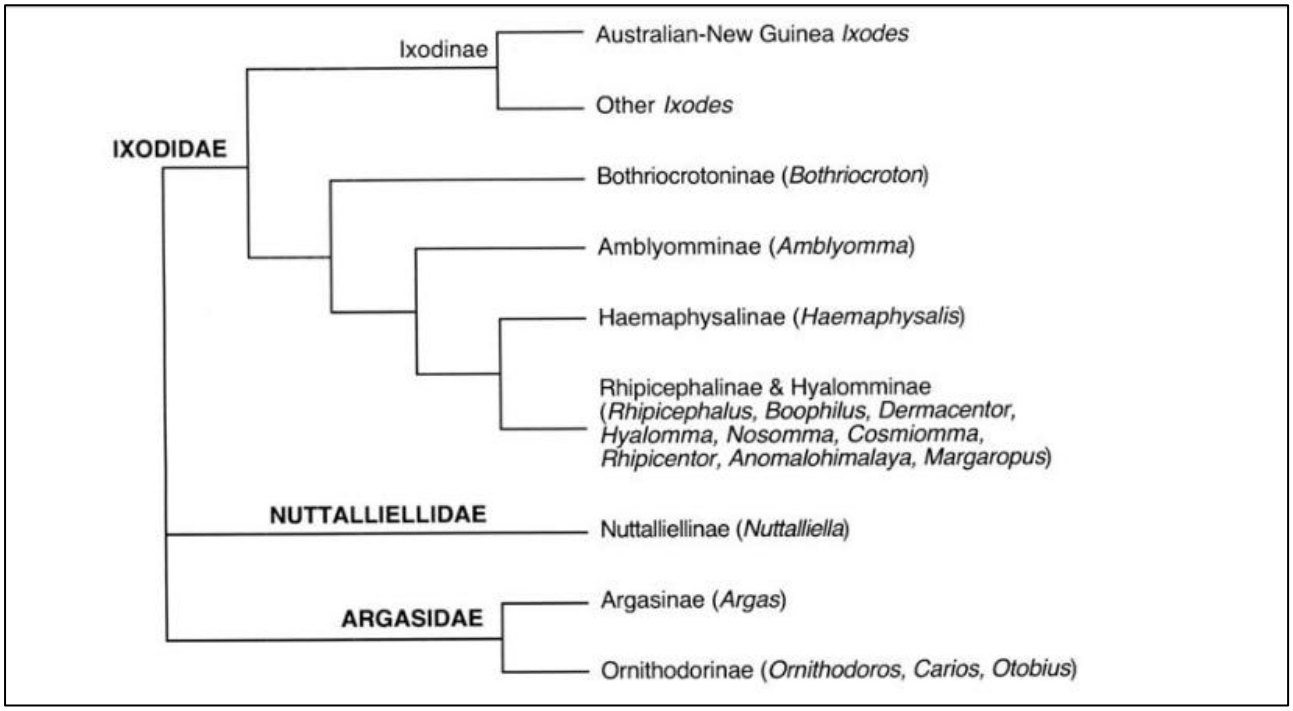


Figure 1.1: The systematic and evolutionary relationship between ticks *(Barker, Murrell, 2004)*.

1. 1. 2. Adaptation to blood-feeding

Hemostasis is a process which results in the arrest of bleeding, when a blood vessel is ruptured. This provides animals with a first line of defense against hematophagous parasites. Hematophagy, the practice of ticks to use blood as a food source, may have evolved during the Cretaceous and Jurassic eras, between145-65 million years ago (Graca-Souza *et al.*, 2006). As a result, ticks have evolved immunomodulatory and anti-hemostatic proteins to circumvent many of the host's defense systems. Therefore, the discovery and classification of these anti-hemostatic components from tick saliva has aided in the understanding of tick evolution regarding adaptation to a blood-feeding environment (Mans *et al.*, 2011; Ribeiro, 1995).

Adaptation of ticks to a blood-feeding behavior was first suggested to be independent to host specificity and rather as a result of ecological factors (Klompen *et al.*, 1996). However, more recent reports suggest that this evolution is independent of environmental changes. Hence, host variety could indeed have played a significant role on tick-vertebrate hemostatic adaptations (Mans *et al.*, 2002).

1. 1. 3. Tick life cycles

The development of specialized life cycles, such as one-host and two/three-host life cycles, are additional adaptations that allow ticks to be very successful parasites (Troughton, Levin, 2007). For ticks with one-host life cycles all stages feed and molt on the same host, whereas for two host life cycles, the immature stages (larvae and nymphs) feed on the same host while adults feed on a second host. In contrast to ticks with three-host life cycles, all three stages occur on separate hosts. Molting on the same host is a benefit to one-host ticks, as the risk of finding a suitable host to continue the life cycle is reduced (Li *et al.*, 2007).On the other hand, advantages of two/three-host life cycles include increased reproduction and the tick's ability to live without food for a longer period of time compared to one-host ticks (Troughton, Levin, 2007).

Ticks have four main developmental stages that include egg, larval, nymphal and adult stages (Figure 1.2). Male ticks copulate with more than one female and

following successful feeding and mating; engorged female ticks drop from the host to lay eggs. The eggs develop within the female tick for some 36-57 days prior to oviposition. Female hard ticks die within days after laying between 1000 to 10 000 eggs, as much of the internal viscera is degraded to provide energy and accommodate the developing clutch of eggs (Vial, 2009). In contrast, most soft ticks can feed numerous times and lay from 20 to 50 eggs after each blood meal. Six-legged larvae hatch from the eggs within two weeks or a few months, depending on environmental factors, such as humidity and temperature (Estrada-Pena, 2008). Once engorged, larvae drop off the host and molt into eight-legged nymphs. The larvae of one-host ticks, such as *R. microplus*, remain on the same host during molting. The following nymphal stage is similar to the adult tick, but lacks a genital opening. The nymph awaits a suitable host, attaches and after engorgement, drops from the host and subsequently molts to an adult tick (Vial, 2009).

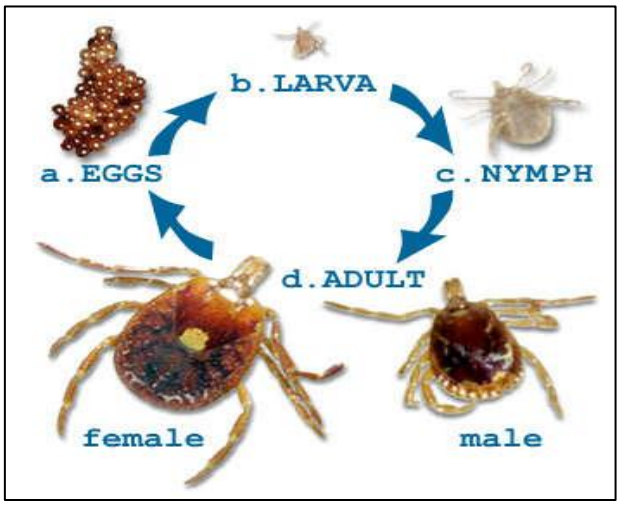


Figure 1.2: Tick developmental stages. *The life cycle of ticks includes four stages: egg, six-legged larva, eight-legged nymph, and adult. The engorged female (d) lays 1000 to 10 000 eggs, and dies. Eggs (a) hatch to larva (b) and the larva engorge and molt into a nymph (c). Nymphs feed and molt to an adult (d). After engorgement the female tick mates with a male before dropping from the host (Figure taken from: http://www.ticktexas.org/ticks/ticks101_develop.htm. [2013, January 4]).*

The lifespan of ticks differs from species to species, with some hard tick species completing their life cycle in a few weeks or months and soft tick species taking on average longer to complete their life cycle, due to having at least two nymphal

stages (Oliver, 1989). In some tick species, such as the American dog tick, *Dermacentor variabilis*, the entire life cycle takes three years to complete (Burg, 2001).

1. 1. 4. Tick feeding

Ticks can ingest a blood meal 5 to 10 times their own body weight, forcing the tick's body to expand (Figure 1. 3). Ixodid ticks are slow-feeding organisms, requiring up to 25 days to feed. The dorsal scutal plate allows them to attach more strongly to the host skin, as they have to remain longer on the host to feed to completion. On the other hand, argasid ticks do not have a scutal plate, requiring only a short period of time on the host to feed successfully. Therefore, they have folded leathery cuticles which allow them to expand rapidly, resulting in shorter feeding times that last 1-2 hours (Mans *et al.*, 2004; Sonenshine, 1991).

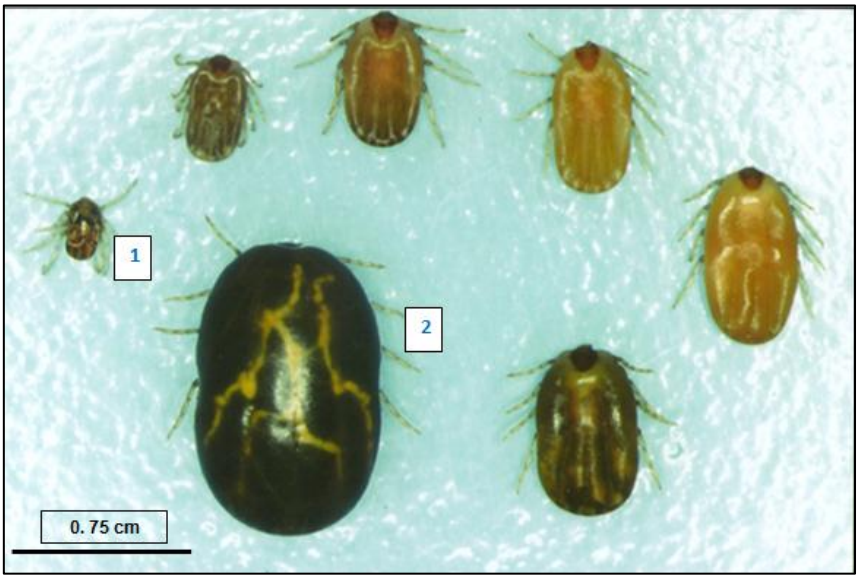


Figure 1.3: *Rhipicephalus microplus* female at various stages of engorgement from unfed (1) to fully engorged (2). *After larvae molt into nymphs they become unfed young adults (1). It takes 22-25 days to become a fully engorged adult tick (2)(Figure taken from: www. daff. gld. gov. au/ [2012, September 25]).*

Prior to becoming obligate parasites, ticks might have fed on lymph (Mans *et al.*, 2004). This is supported by the fact that the first few days of feeding do not necessitate the ingestion of blood but lymphatic fluid, resulting in a faster fluid intake (Donohue *et al.*, 2009; Mans *et al.*, 2012). Ticks feed on the host's blood by penetrating their hypostomes into the skin, creating a feeding pool by rupturing vessels and subsequently ingesting the blood. This is a challenge to the tick as the host responds with mechanisms involved in blood coagulation, immunity and inflammation (Hovius *et al.*, 2008; Mans *et al.*, 2012). However, the saliva of ticks contains a mixture of bioactive compounds, including anticoagulants and immunomodulatory compounds, which allow successful attachment and feeding. These proteins inhibit the hemostatic response of the host maintaining the fluidity of the blood meal (Ribeiro *et al.*, 2005). Some of the anticoagulants found in tick saliva are factor Xa inhibitors (e.g., TAP-tick anticoagulant peptide and Salp14), tissue factor pathway inhibitors (e.g., Ixolaris and Pentalaris) and direct thrombin inhibitors (e.g., microphilin, savignin, ornithodorin, madanin-1 and -2, as well as variegin) (Hovius *et al.*, 2008; Simons *et al.*, 2011). Components that act as immunosuppressants in the tick saliva include complement inhibitors (e.g., Isac and Salp20) and inhibitors of T- and B-cells (e.g., Salp15, Iris and B-cell inhibitory factors) (Brossard, Wikel, 2004; Camargo Mathias *et al.*, 2011; Valenzuela, 2004). The immunomodulatory compounds found in tick saliva are also thought to contribute to successful pathogen transmission by ticks, such as Salp 15 (Hovius *et al.*, 2007). Tick feeding therefore plays a dual role in pathogen transmission, directly by allowing the transmission of pathogens into the host during salivation and indirectly by modulating the host's immune response (Camargo Mathias *et al.*, 2011; Kazimirova, 2008; Nuttall, Labuda, 2004).

1. 1. 5. Tick-borne diseases and direct effects of tick infestation on hosts

Ticks are one of the most common vectors of human and animal diseases (Hovius *et al.*, 2008). Currently, there are more than 70 tick species established as economically important vectors of rickettsial (such as anaplasmosis and heartwater or cowdriosis), protozoan (such as theilerioses and babesiosis), and spirochete (such as borreliosis and relapsing fever) diseases. Along with the transmission of

bacterial diseases, they are also vectors for viruses, such as tick-borne encephalitis and Crimean-Congo hemorrhagic fever (Table 1.1) (Jongejan, Uilenberg, 2004).

Table 1.1: Summary of tick-borne diseases, associated pathogens and their respective tick vectors.

	Disease	Organism	Vector	References
Bacteria	Lyme Disease (Borreliosis)	*Borrelia burgdorferi*	*Ixodes scapularis, I. pacificus, I. ricinus , I. persulcatus, Rhipicephalus microplus*	(Embers, Narasimhan, 2013) (Nadelman, Wormser, 2013)
	Relapsing fever	*Borrelia* spp.	*Ornithodoros* spp.	(Nieto *et al.*) (Teglas *et al.*, 2011)
	Typhus	*Rickettsia* spp.	*Dermacentor andersoni*	(Botelho-Nevers *et al.*, 2012) (Gage *et al.*, 1992)
	Rocky Mountain Spotted Fever	*Rickettsia rickettsii*	*D. variabilis, D. andersoni*	(Wikswo *et al.*, 2008) (Atkinson *et al.*, 2012) (Minniear, Buckingham, 2009)
	Helvetica Spotted fever	*R. helvetica*	*I. ricinus*	(Boretti *et al.*, 2009) (Sprong *et al.*, 2009)
	Ehrlichiosis and anaplasmosis	*Anaplasma phagocytophilum, A. marginale*	*Amblyomm aamericanum, I. scapularis, R. microplus*	(Kalinova *et al.*, 2009) (Kelly *et al.*, 2013) (Zivkovic *et al.*, 2010)
	Tularemia	*Francisella tularensis, A. americanum*	*D. andersoni, D. variabilis*	(Graham *et al.*, 2011)
Viruses	Meningo encephalitis	Flavivirus	*I. scapularis, I. ricinus, I. persulcatus*	(Lindquist, Vapalahti, 2008)
	Colorado tick fever	Coltivirus	*D. andersoni*	(Bratton, Corey, 2005) (Klasco, 2002)
	Crimean-Congo hemorrhagic fever	Nairovirus	*Hyalomma marginatum, R. bursa*	(Leblebicioglu *et al.*, 2012) (Ergonul, 2012)
Protozoa	Babesiosis	*Babesia microti, B. equi, B. bigemia, B. bovis*	*I. scapularis, I. pacificus, R. microplus*	(Suarez, Noh, 2011) (Graham *et al.*, 2011)
	Cytaux zoonosis	*C. felis*	*D. variabilis*	(Reichard *et al.*, 2010)

A further concern relating to ticks is that they transmit pathogens that are responsible for causing zoonotic diseases (Homer *et al.*, 2000). However, the knowledge of factors influencing the maintenance and spread of zoonotic diseases is currently limited (Stephen *et al.*, 2004). Tick-borne zoonotic diseases are caused by various microorganisms such as *Babesia* spp., *Borrelia* spp., *Rickettsia* spp., *Ehrlichia* spp.,

Francisellatularensis, *Coxiellaburnetii* and tick-borne encephalitis virus. Lyme disease, human babesiosis and human ehrlichiosis are the most well-known tick-borne zoonotic diseases (Graham *et al.*, 2011). Another example of a zoonotic disease is dermatophilosis (mud fever), a skin disease caused by a tick-borne fungus *Dermatophilus congolensis*. It was previously reported in South Africa and Britain to affect both humans and livestock (Harwood *et al.*, 1979). To date, the number of cases of zoonotic incidences has increased dramatically over the last few years posing a major concern for global health (Perez de Leon *et al.*, 2010).

Besides the pathogens, ticks transmit various compounds found in tick saliva which can lead to toxicosis. The severe types of toxicosis include tick paralysis (neurotropic toxin), sweating sickness (dermotropic toxin) and *Rhipicephalus appendiculatus* toxicosis (leukotropic toxin) (De la Fuente *et al.*, 2007b; Sonenshine, 1991). Furthermore, tick infestation results in hindered cattle growth as the loss of blood leads to a reduction in weight gain and milk production (De la Fuente *et al.*, 2007b).

Economic impact of ticks

Together with the repertoire of pathogens ticks transmit, tick infestation results in devastating effects on the economy (Hovius *et al.*, 2008). Annual economic losses as a result of tick infestation are estimated to be billions of dollars globally (Andreotti *et al.*, 2011).

In South Africa, 3.3 million US$ of the agricultural sector's gross domestic product (GDP) was generated by the livestock industry which included hides, milk and beef (The State of Food and Agriculture, FAO, 2005). Additionally the cattle industry provides financial support to 2.6 million people (Department of Agriculture, Fisheries and Forestry, DAAF, 2011). Also approximately 69% of agricultural land is used for pastures illustrating the importance of the livestock industry. Consequently, it is evident that disease outbreaks would have resulted in a tremendous cost. Therefore, it is vital to study and understand tick physiology to be able to control tick infestation, and to prevent economic losses and the transmission of zoonotic diseases

1. 1. 6. *Rhipicephalus microplus*

i) Geographical distribution

A tick's adaptability to its living (e.g., host, competitors, parasites, predators) and non-living environment (e.g., climate and weather), determines its geographic distribution (Gomez-Diaz *et al.*, 2012; Gray *et al.*, 2010). For example, multivariate analysis of tick distribution in Africa showed that climate change will allow an increase in the distribution of the ticks into areas previously unsuitable, such as the Northern and Eastern Cape Provinces of South Africa, Botswana and Zambia (Olwoch *et al.*, 2008).

Rhipicephalus microplus is a one-host tick of cattle found throughout tropical and subtropical regions of the world, which include the southern parts of Asia, Madagascar, as well as southern and eastern Africa (Estrada-Peña *et al.*, 2006; Estrada-Peña *et al.*, 2012a). *Rhipicephalus microplus* initially only occurred in tropical and subtropical areas, but the tick's adaptability and changing environmental conditions has led to an increased distribution of *R. microplus*. Recent survey data indicate that *R. microplus* is still extending its distribution range, invading previously non-occupied areas, such as northern Tanzania (Lynen *et al.*, 2008) and West Africa (Madder *et al.*, 2011). *R. microplus* is also abundantly found throughout South Africa (Terkawi *et al.*, 2011). Initially, *R. microplus* was introduced to the coastal areas of South Africa by cattle imported form Madagascar during the late 1800s. Reports from the early 1900s indicate that *R. microplus* replaced endemic *R. decolaratus* in the Cape Province (Nyangiwe *et al.*, 2011). Figure 1.4 indicates the distribution of *R. microplus* and the indigenous species *R. decoloratus.*

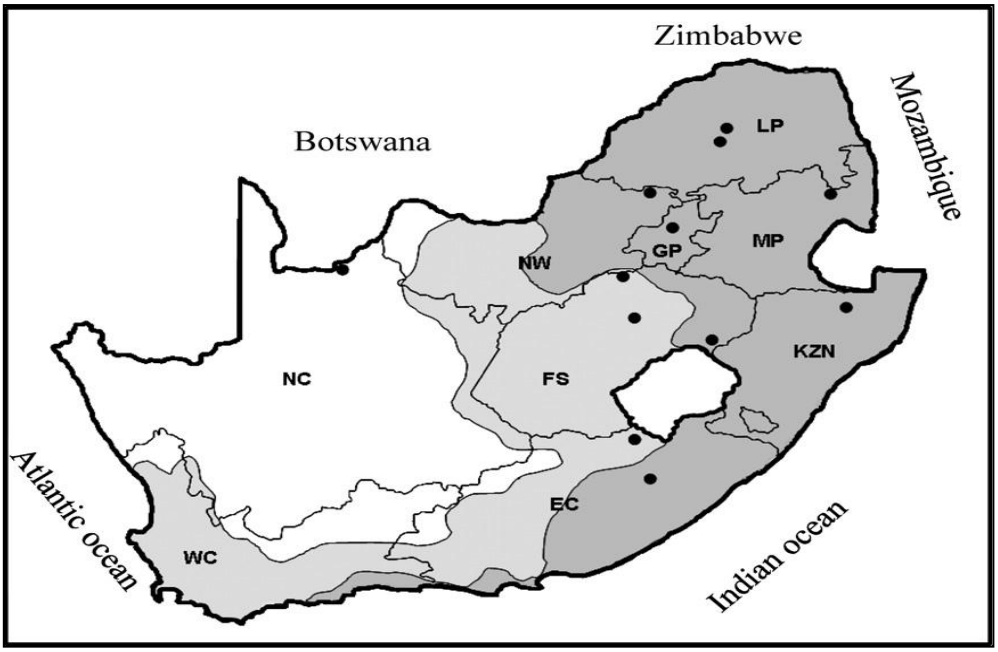


Figure 1.4: Distribution of *R. microplus* in South Africa (Terkawi *et al.*, 2011)**.** Dark gray areas represent the occurrence of both *R. microplus* and *R. decoloratus* ticks and light gray areas represent the occurrence of only *R. decoloratus*. Abbreviations correspond to Limpopo Province (LP), Mpumalanga (MP), Gauteng Province (GP), North-West (NW), Kwazulu-Natal (KZN), Free State (FS), Northern Cape (NC), Eastern Cape (EC) and Western Cape (WC).

ii) Life cycle

Rhipicephalus microplus is a one-host tick, predominantly infesting cattle from the *Bos* genus. (Sonenshine, 1991). *Rhipicephalus. microplus* completes its life cycle in approximately 3 weeks and an adult female *R. microplus* can lay up to 3000 eggs. Therefore, the life cycle of consecutive generations can be repeated several times a year (Walker *et al.*, 2003). With such high fecundity, this tick species results in heavy seasonal tick burdens, drastically affecting the livestock industry. During the hot seasons, *R. microplus* larvae can survive for up to 4 months without feeding. However, in the colder seasons, they may survive without food for up to six months (Barrero *et al.*, 2011).

iii) Diseases

Rhipicephalus microplus is predominantly a parasite of cattle, however infestations may occur on horses, sheep, deer, water buffalo, goats, dogs, marsupials, cats and pigs (Barker, Murrell, 2002; Barre, Uilenberg, 2010). The clinical signs associated with tick infestation on cattle hosts include reduced milk production, anemia and weakness (Domingos *et al.*, 2013). Even though *R. microplus* can cause considerable damage simply by the large amount of ticks that infest an animal (e.g., hide damage, increased secondary infection, blood loss), its greatest threat to the livestock industry is ascribed to its ability to transmit the causative agents of babesiosis, anaplasmosis and borreliosis (Table 1.1) (Madder *et al.*, 2007; Tonnesen *et al.*, 2004). Anemia develops as a result of anaplasmosis, which is initiated by extensive erythrolysis (lysis of red blood cells) and anti-erythrocytic antibodies (Graham *et al.*, 2011; Zivkovic *et al.*, 2010). Further symptoms include weakness, decreased milk production, loss of appetite, rapid weight loss, fever of up to 41ºC and life-threatening anoxia that can potentially also lead to abortions. The use of antibiotic treatment is generally only effective if administered during the early developmental stage of the parasite. Therefore, prevention by vaccination rather than treatment is important (De Castro, 1997; Kocan *et al.*, 2003).

One of the most economically important pathogens transmitted by *R. microplus* is the tick-born protozoan *Babesia bovis* that causes Asiatic bovine babesiosis. *Babesia bovis* is a member of the Apicomplexa and is transmitted transovarially from female ticks to the eggs (Hasle *et al.*, 2010). When an infected larva feeds on a host, it releases the parasites into the host's blood circulation and the inter-erythrocytic life cycle starts. Infected red blood cells are destroyed upon parasite release leading to hemolytic anemia. The parasite is then transmitted from the infected host to uninfected ticks, completing the life cycle (Brown, 2001; Suarez, Noh, 2011).

Babesiosis in cattle is regarded globally as a devastating arthropod-borne disease (Bock *et al.*, 2004). Survey data of the last decade from Brazil have shown that eighty percent of more than 300 cattle fever outbreaks were due to *B. bovis* (Bock *et al*., 2004). Resulting economic losses of more than 2 billion dollars in Brazil alone

were reported (Andreotti *et al.*, 2011). As *R. microplus* is a vector for both the protozoan *Babesia bovis*, as well as the rickettsia, *Anaplasma marginale* (responsible for anaplasmosis), it is apparent that the control of *R. microplus* is of tremendous importance to the cattle industry.

iv) Genome studies of R. microplus

The only tick genomic data available up to date is for *I. scapularis*, which was found to have a genome size of $2.1x10^3$ mega bases. Data obtained from expressed sequence tags (ESTs), cDNA library-, bacterial artificial chromosome (BAC) library- and whole genome shotgun sequencing was combined to assemble the draft *I. scapularis* genome (Hill, Wikel, 2005; Pagel Van Zee *et al.*, 2007). In comparison, the size of the *R. microplus* genome was estimated to be $7.1x10^9$ base pairs containing approximately 60% repetitive sequences (Ullmann *et al.*, 2005). However, the main obstacle in gaining more information about tick genomes is their large size and repetitive nature (Jensen *et al.*, 2007; Jongejan *et al.*, 2007).

Wang *et al*. constructed a *R. microplus* expressed sequence tag (EST) database which contained 13, 643 unique transcripts that were assembled from over 42, 000 ESTs. From various tick tissues, strains and life stages a comparative study showed that for more than 60% of the *R. microplus* genome shares no sequence similarity other organism´s genomes (Wang *et al.*, 2007).

Bellgard *et al.* (2012) formed a consortium to begin the sequencing of the southern cattle tick genome and more than 1.86 Gbp from gene-enriched regions of *R. microplus* have been assembled. A comparison between *R. microplus* from Australia and the data contained in the BmiGI database showed that more than 9000 transcripts are unique to the Australian dataset (Estrada-Peña *et al.*, 2012b). An explanation for the high number of unique transcripts was published shortly thereafter, showing that the species in Australia presumed to be *R. microplus* was actually found to be a different species, now called *R. australis* (Estrada-Peña *et al*., 2012b).

1. 1. 7. Current tick control methods

i. Natural control of ticks

One of the main reasons why ticks are very successful ectoparasites is that they have few natural enemies, such as the Oxpecker bird species *Buphagus africanus* and *B. erythrorhynchus* which are found only on the African continent (Nunn *et al.*, 2011). Fungal species such as *Beauveria bassiana* and *Metarhizium anisopliae* have also been found to reduce tick infestation (Fernandes, Bittencourt, 2008; Fernandes *et al.*, 2012). Plant extracts have also been used for the control of ticks. An extract of boiled *Tephrosia vogelii* leaves (locally found in sub-Saharan Africa and Kenya) has been shown to effectively protect livestock against tick infestation (Kalume *et al.*, 2012). The use of natural control measures of ticks is associated with lower environmental and food contamination, a reduction in acaricide resistance development and toxicity to animals and humans. However, the implementation of these control measures is not easy in practice, as there are many challenges to be overcome including the evaluation of their efficacy, finding ways to make them suitable for large scale use, optimal implementation, as well as determining their stability for long-term use and the possibility of non-target effects (Fernandes *et al.*, 2012; Willadsen, 2006b).

ii. Acaricides

The most commonly used chemical acaricides for the control of ectoparasites are cyclodiene organochlorines (Lindane), organophosphates (Diazinon and Fenthion), carbamates (Carbaryl), formamidines (Amitraz), pyrethroids (Permethrin, Deltamethrin and Flumethrin) and phenylpyrazoles (Fipronil and Pyriprole) (Beugnet, Franc, 2012; Willadsen, 2004).

Although the use of acaricides is only partially effective (due to the development of acaricide resistance), it also results in contamination of milk and meat products as well as the environment (De la Fuente *et al.*, 2007; De la Fuente, Kocan, 2003; Willadsen, 2004). Moreover, the development of new acaricides is a lengthy

expensive process and alternative approaches need to be developed for the control of ticks (Fernandes *et al.*, 2012; Guerrero *et al.*, 2012b)

iii. Vaccine development

The current research focus is on the development of anti-tick vaccines, promising an alternative to the control tick infestations (Zintl *et al.*, 2003). A vaccine usually consists of attenuated, killed organisms or their derived proteins to enhance the host immune response to that particular organism (Lawrence, 2005). Once the host's immune system is stimulated in response to vaccination, the foreign organism can be recognized and destroyed. Preventing tick infestations through vaccines also have the added advantages of reducing environmental contamination resulting from acaricide usage and limiting the development of resistance (Willadsen, 2004). Development of an anti-tick vaccine is also seen as a cost effective strategy compared to the usage of acaricides that require frequent re-application (De la Fuente *et al.*, 2007). Furthermore, a vaccine that combines multiple antigens might confer protection against a wider range of tick species (De la Fuente *et al.*, 2007). To date, three antigen types have been described for anti-tick vaccines described in the following sections.

Antigen types

The first type, exposed antigens, are exposed directly to the host immune system during feeding (Nuttall *et al.*, 2006). Exposed antigens such as cystatin, haemangin, immunoglobulin binding protein (IGBMC) and histamine binding protein (HBP) (Table 1.2) are mostly components secreted by the tick salivary glands present in the feeding pool. Using exposed antigens in vaccine development has the advantage that regular booster vaccinations are not necessary, as the repeated exposure to the host will continuously boost the immune system, limiting the need for additional vaccinations (Gourlay *et al.*, 2009; Willadsen, 2004; Willadsen, 2006a; Willadsen, 2006b). However, the immune response induced by these exposed antigens can be insufficient to protect against tick infestation due to co-evolution of the host and the tick (Almazán *et al.*, 2012; Nuttall *et al.*, 2006). It is hypothesized that the latter would allow ticks to become resistant against the host's immune responses (Brake, Perez

De Leon, 2012). Furthermore, vaccines that target only one protein of a large gene family (e.g., proteases) are suggested to be unsuccessful, as the function of the targeted protein could simply be replaced by any other member(s) of the family (Guerrero *et al.*, 2012a). Thus, a vaccine containing only exposed antigens will not necessarily be effective in maintaining a protective response. To date, a number of targets have been identified that are exposed and tested in trails (Table 1.2).

The second antigen type is concealed antigens, which are not exposed to the host immune system during feeding such as antigens located in the midgut and ovaries. Concealed vaccines are usually more immunogenic as they are extracted from tick tissues and inoculated artificially via vaccination into an animal (Nuttall *et al.*, 2006). After vaccinating with such an antigen, the tick is thus exposed to the entire activated host immune response, including antibodies, the complement system, cytokines and activated leukocytes that are taken up during feeding. The possible interaction between host immunoglobulins and the complement system in the hemolymph (and/or gut) of the tick, makes concealed antigens promising targets, as this interaction could result in the rupture of the tick gut similar to that described for the Bm86 antigen (see section 1.1.8) (Guerrero *et al.*, 2012b). It is assumed that a vaccine derived from a concealed antigen will be viable especially when the antigen is part of a organism-specific protein (Guerrero *et al.*, 2012b). The effect of a vaccine could also be enhanced if the target antigen is involved in vital metabolic pathway(s), refer to Table 1.2 (Nuttall *et al.*, 2006).

Consequently a protein possessing the advantages of both exposed and concealed antigens is likely to enhance the efficacy of a vaccine in terms of controlling ticks, as well as the transmission of pathogens (Guerrero *et al.*, 2012b). Vaccines based on proteins that are conserved and/or expressed throughout the tick life cycle have also been proven to be effective against both adult and immature stages (Nuttall *et al.*, 2006). Therefore, the use of highly conserved protective antigens is regarded as very promising in terms of controlling multiple tick species in a wide geographic range (Guerrero *et al.*, 2012b).

In addition to being effective against the tick species of interest, vaccines derived from any antigen should ultimately meet the requirements of an optimal vaccine. These requirements include: activity against various tick species and life stages,

long-lasting immunity, prevention of tick attachment, reduction of disease transmission and being cost effective to implement (Nuttall *et al.*, 2006; Piesman, Eisen, 2008).

Limitations of vaccine development against ticks and tick-borne pathogens

The identification of candidate antigens for vaccine development remains problematic (De la Fuente, Kocan, 2006). Annotation and assembly of the *R. microplus* genome, which is approximately three times the size of the human genome, is one of the main limiting steps (De la Fuente, Kocan, 2006).

Although many promising vaccine candidates have been proposed prior to the availability of genomic/ transcriptomic data, only few antigens have been tested in trials (Decrem *et al.*, 2008; Francischetti *et al.*, 2003). Table 1.2 gives a list of all the described antigens to date. The control of tick infestation through vaccination alone might not be sufficient to prevent problems associated with ticks in the short run, but it may reduce the usage of acaricides and therefore reduce the development of acaricide-resistance. Finally, problems with field evaluation and commercialization, which are some of the fundamental components of vaccine development, should also be considered to permit successful tick and tick-borne pathogen control (De la Fuente *et al.*, 2007; De la Fuente, Kocan, 2003).

Table 1.2: Ixodid exposed, concealed and dual antigens described to date (Botha, 2013). *Vaccination efficacy is indicated by: (n/d) not determined; (-) no significant effect; (+)>25% efficacy; (++)25 – 50% efficacy; (+++)50 – 75% efficacy; (++++)75 – 100% efficacy.*

	Tick species	Characteristics/Putative function	Vaccination efficacy
Exposed antigens			
Calreticulin	*Amblyomm aamericanum, Dermacentor variabilis, R. microplus.*	Endoplasmic reticulum (ER) calcium binding protein. Lacks ER retention (KDEL) sequence. Role in feeding. Low immunogenicity in cattle.	**n/d** (Jaworski et al.,1995) (Ferreira et al.,2002b)
Immunoglobulin binding protein(IGBMC).	*R. appendiculatus*	Male-specific salivary protein with an immunomodulatory function. Slight effect on female feeding performance on vaccinated guinea pigs.	**n/d** (Wang and Nuttall, 1999) (Wang et al., 1998)
Histamine binding protein (HBP).	*R. appendiculatus*	Lipocalin-like fold. Suppresses host inflammation. Diverse reaction in guinea pig model.	**n/d** (Paesen et al., 1999)
P29	*Haemaphysalis longicornis*	Salivary gland putative extracellular matrix protein. Collagen like. Effective against all tick stages.	**++** (Mulenga et al., 1999)
HL34	*H. longicornis*	Salivary protein. Adverse effect on tick feeding on rabbits.	**+** (Tsuda et al., 2001)
HLMP1	*H. longicornis*	Salivary proteins with fibrinogenolytic and gelatinase activities and function in blood feeding. Contains a signal sequence, metalloprotease characteristic zinc-binding motif, and cysteine rich	**+** (Imamura et al., 2009)

		region.	
RIM36	*R. appendiculatus*	Cement protein. Antigenic but not protective in cattle.	**+** (Bishop et al., 2002) (Havlíková et al., 2009)
Tick histamine release factor(tHRF).	*I. scapularis*	Secreted in tick saliva. Stimulates histamine release. Silencing (RNAi): impaired feeding and decreased *Borrelia burgdorferi* infection in mice.	**n/d** (Dai et al., 2010)
HL-p36	*H. longicornis*	Immunosuppressor. Inhibited proliferation of mitogen stimulated immune cells; associated decrease in interleukin-2 messenger RNA (mRNA) levels. rHL-p36-inoculated mice: lower proliferative response of splenocytes in vivo and down regulation of immunomodulating genes.	**n/d** (Konnai et al., 2008)
Haemangin (Kunitz-type inhibitor)	*H. longicornis*	Disrupted angiogenesis and wound healing (induces apoptosis). Inactivated trypsin, chymotrypsin and plasmin (anti-proteolytic). RNAi: Decreased engorgement success and failed angiogenesis disruption.	**n/d** (Islam et al., 2009)
Evasin-3	*A. variegatum* *R. appendiculatus* *D. reticulatus*	7 kilo Dalton (kDa) Salivary protein. Binds neutrophil chemoattractants, CXCL8 and CXCL1. Prevents influx of neutrophils to feeding site.	**n/d** (Vancova et al., 2010b)
Sialostatin L/L2	*I. scapularis*	Salivary C1-type cysteine protease inhibitors. Facilitates blood feeding and transmission of *B. burgdorferi.*	**++** (Kotsyfakis et al., 2006) (Kotsyfakis et al., 2010)
Cystatin (HISC-1)	*H. longicornis*	Protease inhibitors of Papain-like Cysteine proteases. Expressed in type II aciniof salivary glands during early blood feeding.	**n/d** (Yamaji et al., 2009)
Longistatin	*H. longicornis*	Anticoagulant, activates plasminogen. EF-hand Ca2+ binding domains. Silencing diminished blood pool formation, feeding.	**n/d** (Anisuzzaman et al., 2011)
Hq05	*H. qinghaiensis*	Novel gene with ORF of 540 bp. Expressed in salivary glands of nymphal and adult stages. Significant effect on oviposition.	**++** (Gao et al., 2009)
HqCRT	*H. qinghaiensis*	ORF of 1,233 bp and identified as calreticulin. Ubiquitously expressed in different tissues and life stages.	**+** (Gao et al., 2008b)
HqTnT	*H. qinghaiensis*	Troponin-T is involved in Ca2+-sensitive molecular switching of muscle contraction.	**-** (Gao et al., 2008a)
IRACs	*I. ricinus*	Paralogous anti-complement proteins co-expressed in salivary glands.	**-** (Gillet et al., 2009)
IRIS	*I. ricinus*	Ubiquitous serpin family of proteins with high affinity for human leukocyte elastase. Interferes with hemostasis and immune response of host. Also inhibits lymphocyte proliferation and the secretion of interferon- and/or tumornecrosis factor-, platelet adhesion, coagulation and fibrinolysis.	**++** (Prevot et al., 2009)
Metis 1	*I. ricinus*	Metalloprotease putatively involved in tissue alterations through digestion of its structural components, thereby interfering with homeostasis.	**++** (Decrem et al., 2008)
Dual antigen			
64P (64TRP)	*R. appendiculatus* *I. ricinus* *R. sanguineus*	Cement protein. Similar composition to vertebrate keratin and collagen. Dual action by cross reacting with concealed (midgut), and exposed (salivary) antigens. Potentially "universal" affecting many different tick species. Exhibited average 35% mortality of *I. ricinus* immunized rabbits.	**+++** (Trimnell*et al.*, 2002) (Trimnell*et al.*, 2005) (Havlíková*et al.*, 2009)
Concealed antigens			
Bm86	*R. microplus*	Gut cell surface protein (glycoprotein). Commercial vaccine: Gavac®. Possible function in the regulation of gut digest cell growth.	**+++ (recombinant)** **++++ (native).** (De la Fuente and Kocan, 2003) (Willadsen, 2004) (Tellam et al., 1992) (Willadsen et al., 1989)
Ra86	*R. appendiculatus*	Two homologues identified (654 and 693 amino acids with 80% amino acid identity. The 654 amino acid homologue is transcriptionally dominant. This represents an example of tick allelic exclusion (different alleles/variants at single locus).	**-** (Kamau et al., 2011)
Ba86	*R. annulatus*	Immunization of cattle with rBa86: reduced tick infestation, weight, oviposition, and hatching for *R. microplus* and *B. annulatus*. Better efficacy for *B.*	**++++** (Canales et al., 2009a)

		annulatus – possibly due to specific antigenic epitopes.	
Bm86	*A. variegatum*	Gut cell surface protein (glycoprotein).	**-** (de Vos et al., 2001)
Bm86	*Hy. anatolicum*		**++** (de Vos et al., 2001)
Bm86	*Hy. dromedarii*		**++++** (de Vos et al., 2001)
Bm86	*R. decoloratus*		**+++** (de Vos et al., 2001; Odongo et al., 2007)
Bm86	*R. sanguineus*		**++** (Perez-Perez et al., 2010)
Haa86	*Hy. anatolicum*		**+++** (Azhaniahambi et al., 2009)
Bm91	*R. microplus*	Carboxydipeptidase. Supposed to increase efficacy of Bm86.	**++** (Willadsen et al., 1996)
Bm95	*R. microplus*	Gut protein of unknown function. Bm86 allelic variant cloned from Argentinean A strain (64 kDa) used in Australian vaccine, TickGARD® Plus. Protects cattle against infestations by *R. microplus* strains from different geographical areas.	**++++** (De la Fuente and Kocan, 2003) (Willadsen, 2004) (Kumar et al., 2009) (Garcia-Garcia et al., 2000)
Bm95	*R. haemaphysaloides*	Protected Bosindicus cattle against R. haemaphysaloideslarval, nymphal and adult infestation.	**++++** (Sugumar et al., 2011)
Vitellin	*R. microplus*	Abundant egg protein. Product of vitellogenin. Ovine vaccination with native protein: increased mortality, reduced weight and oviposition. 45% Engorged females displayed damaged, red phenotype.	**++** (Tellam et al., 2002)
VTDCE (vitellindegrading cysteine endopeptidase)	*R. microplus*	Vitellin hydrolysis in eggs during tick development. Enzyme activity in ovary, gut, fat body, salivary gland and haemolymph. Within female gut: localized to areas of protein synthesis and trafficking with haemolymph.	**+** (Seixas et al., 2008) (Seixas et al., 2010)
BmPRM(paramyosin).	*R. microplus*	Multi-functional, conserved invertebrate muscle protein. rBmPRM binds IgG and Collagen. Roles in immunomodulation.	**n/d** (Ferreira et al., 2002a)
HLS1/HLS2	*H. longicornis*	Serine protease inhibitor (Serpin). Rabbit vaccination with rHLS1: 43.9% mortality of nymphs.	**+/++** (Sugino et al., 2003) (Imamura et al., 2005)
Voraxin	*A. hebraeum*	Male engorgement factor; impairs feeding (<74%) of females in a rabbit model; salivary gland degeneration, partial ovary development.	**n/d** (Weiss and Kaufman, 2004)
Voraxina	*R. appendiculatus*	Expression induced by blood feeding. Vaccination of rabbits: elicited humoral and protective immunity against female ticks and reduced weight.	**+++** (Yamadaa et al., 2009)
P27/30 (troponin-1 like protein).	*H. longicornis*	Troponin I-like protein. Impairs tick feeding on immunized mice. Low larval attachment: 31.1%.	**n/d** (Myung-Jo, 2005)
BmTI	*R. microplus*	Trypsin and serine proteinase inhibitor. Reduction in population (69.7%) and egg weight (71.3%).Functions in larval attachment and feeding.	**+++** (Andreotti et al., 2002)
4F8	*I. scapularis*	Nucleotidase.	**++** (Almazan et al., 2005a) (Almazan et al., 2005b)
Subolesin (4D8).	*I. scapularis*	Putative function in gene regulation, tick feeding and reproduction. Detailed section follows.	**+++** (Almazan et al., 2003c) (Almazan et al., 2003b) (Almazan et al., 2005a) (Almazan et al., 2005b) (Almazan, 2005)
4E6	*I. scapularis*	Unknown function. Glutamine and Alanine rich protein.	**-** (Almazan et al., 2005a) (Almazan et al., 2005b)
(Elongation factor 1a)	*R. microplus* *R. microplus* *B. annulatus*	Involved in feeding and reproduction, function still to be proven. RNAi: increased mortality and reduced oviposition. Could not produce sufficient rEF-1a for vaccination.	**n/d** (De la Fuente et al., 2008c) (Almazan et al., 2010)
Serine proteinase (HISP, HISP2, HISP3).	*H. longicornis*	Digestion (blood feeding), immune defense. Transcribed in the midgut, lumen. Silencing: Reduced engorgement weight.	**n/d** (Miyoshi et al., 2008)
Varisin (defensin)	*Dermacentor variabilis* *I. scapularis* *A. americanum* *A. hebraeum* *R. microplus*	Small cationic peptides that provide immunity against gram +/- organisms and Babesia spp. Silencing varisin – reduced *A. maginale* numbers.	**n/d** (Johns et al., 2001) (Kocan et al., 2008)
5'-Nucleotidase	*R. microplus*	Ectoenzyme bound by a glycosyl phosphatidylinositol (GPI) anchor to plasma membrane. Degrades	**-** (Hope et al., 2010)

		nucleotide mono, di-and triphosphates for purine salvage. Located in the malphigian tubules. Sheep (challenged with freshly moulted adults) and cattle (challenged with larvae) vaccinated with r5'-nucleotidase. Significant difference in antibody titre between cattle and sheep. No significant protection.	
ATAQ	*R. microplus* *A. variegatum* *D. reticulatus* *H. elliptica* *H. longicornis* *H. anatolicum anatolicum* *H. marginatum* *I. ricinus* *I. scapularis* *R. appendiculatus* *R. evertsievertsi*	Orthologue of Bm86 with conserved signature peptide YFNATAQRCYH ; signal peptide; epidermal growth factor (EGF) -like domains. Silencing provides weak phenotype.	**n/d** (Nijhof A.M. et al., 2010)
Rhipilin-1	*R. microplus*	Kunitz-type anticoagulant. Silencing: significant decrease in attachment rate and body weight.	**n/d** (Gao et al., 2011)
Vitellogenins (Vg1, Vg2, HIVg-1, HIVg-2, HIVg-3).	*D. variabilis* *H. longicornis*	Major yolk protein (vitellin) precursor (phosphoglyco protein). N-terminal lipid binding domain, von Willebrand factor. RNAi results shows importance for egg development and oviposition.	**n/d** (Sullivan et al., 1999) (Boldbaatar et al., 2010b) (Khalil et al., 2011)
I. scapularis antifreeze glycoprotein (IAFGP).	*I. scapularis*	Binds and restricts ice crystal growth enhancing tick survival in extreme cold. Induced by *Anaplasma phagocytophilum* infection	**n/d** (Neelakanta et al., 2010)
Amblyomin-X (Kunitz type inhibitor).	*A. cajennesse*	Kunitz type protease inhibitor. Inhibitor of Factor Xa (blood coagulation).	**n/d** (Batista et al., 2008)
Small heat-shock protein (sHSPII).	*B. annulatus*	Alpha crystalline, small heat shock proteins superfamily. Highly immunogenic. Confers thermal protection to other proteins (molecular chaperones).	**n/d** (Shahein et al., 2010)
Class B scavenger receptor (CD36)	*H. longicornis*	Cell surface proteins. Recognition of microbial surface antigens. Silencing led to reduction in engorged female weight.	**n/d** (Aung et al., 2011)
Lipocalins (e.g. LIPERs).	*I. ricinus* *I. persulcatus* *R. appendiculatus*	Low molecular weight. Immunomodulatory functions: cellular homeostasis and inhibition of platelet aggregation and complement. Vaccination of mice resulted in delayed engorgement and reduced weight.	**-** (Konnai et al., 2011)
Legumain (aspariginyl endopeptidases) HILgm1 and HILgm2. IrAE	*H. longicornis* *I. ricinus*	Midgut specific. Haemoglobin digestion cascade. Silencing (RNAi) of HILgm1 and HILgm2: Failure to reach repletion, reduced engorgement weight, damage to midgut tissue, reduced oviposition, aberrant egg development, and feeding- subsequent cellular remodeling.	**n/d** (Alim et al., 2009) (Sojka et al., 2007)
Kynurenine aminotransferase (HIKAT).	*H. longicornis*	Enzyme of tryptophan pathway (kynurenine pathway). Transcribed in all tissues. In vitro inhibition of *Babesia caballi* growth.	**n/d** (Battsetseg et al., 2009)
GATA factor	*H. longicornis*	Transcribed in midgut and fat body. Transcriptional activator of Vg gene. Knockdown: Inhibition of Vg and significantly reduced oviposition.	**n/d** (Boldbaatar et al., 2010a)
Aquaporin (IrAQP1), (RsAQP1).	*R. sanguineus* *I. ricinus*	Water channel; expressed in gut and salivary glands. Reduced blood ingestion and engorgement weight. Affects gut to saliva water flux and haemolymphosmolarity.	**n/d** (Campbell et al., 2010)
HlYKt6. Soluble Nethylmaleimide sensitive fusion protein attachment protein receptors (SNAREs). Synaptobrevin (VAMP2).	*H. longicornis* *A. americanum*	Functions in signaling, exocytosis. Silencing (RNAi): Significant decrease in engorgement weight and increase in mortality. Decreased salivation and salivation protein repertoire. Knockdown: decrease in engorgement rate and body weight.	**n/d** (Gong et al., 2009)
Ferritin 2 (IrFER2 and RmFER2).	*I. ricinus* *R. microplus* *R. annulatus*	Iron storage/homeostasis proteins. No vertebrate orthologs. Vaccination: Reduced weight and fertility and increased mortality – significantly better results in *I. ricinus*.	**+++** (Hajduseka et al., 2010)
Leucine aminopeptidase (HlLAP).	*H. longicornis*	Localized in ovarian cells. Regulatory role in vitellogenesis. HlLAP knockdown: Oocytes with abnormal morphology, reduced oviposition.	**n/d** (Hatta et al., 2010)
Hemelipoglycoprotein (HeLp).	*D. marginatus* *R. microplus*	Carrier protein of heme in hemolymph (290 kDa): Able to agglutinate red blood cells. Complexes with fibrinogen related proteins.	**n/d** (Dupejova et al., 2011)
Mimotopes	*R. microplus*	Cattle vaccination: causes hemorrhagic events in the	**+**

generated by Phage Display.		gastrointestinal and–reproductive tracts.	(Prudencio et al., 2010a)
Chitinase (CHT1)	*H. longicornis*	Degrades older chitin during molting. Induced by ecdysteroids. Vaccination: Longer feeding and molting periods (76.7% increased molting rate) and reduced egg weight.	**+** (Fujisaki and You, 2009)
BMA7	*R. microplus*	Mucin-like membrane glycoprotein (63 kDa). Similarity to vertebrate mucins and unknown function.	**+** (McKenna et al., 1998)
BYC (Boophilus Yolk Cathepsin)	*R. microplus*	Aspartic proteinase involved in embryogenesis. rBYC used to vaccinate cattle: overall protection of 25.24%. Dose-dependent decrease in oviposition.	**+** (native protein) **++** (recombinant protein)
RAS-3 (Rhipicephalus Appendiculatus Serpin 3), RAS-4 and RIM36.	*R. appendiculatus*	Serine protease inhibitors with important roles in blood coagulation, fibrinolysis and complement activation. Assessed as anti-tick cocktail vaccine in cattle. Resulted in tick mortality rate of 39.5%, and 48.5% in *Theileria parva* infected female ticks.	**++** (Imamura et al., 2008)
Ubiquitin	*R. annulatus*	Ubiquitin (GenBank Accession number: AA257892) was used for RNAi and vaccination experiments. Inconclusive results and low immunogenicity.	**+/+++** (Almazan et al., 2010)
Ubiquitin	*R. microplus*		

1.1.8. Bm86

The only commercially available tick vaccine is currently sold in South America as GAVAC^{TM} and based on Bm86, a membrane-bound concealed antigen found in the gut of *R. microplus* (De la Fuente, Kocan, 2003). Although it is known that Bm86 is located on the surface of gut digest cells, its function remains unknown (Gough, Kemp, 1993; Liao *et al.*, 2007). Determining its function would provide invaluable insight into the mechanism of the protein, paving the way to an improved Bm86-based vaccine (Kamau *et al.*, 2011).

Rand *et al*. (1989) characterized a fragment of cDNA with a 1982 bp open reading frame encoding the Bm86 protein (Figure 1.5). They found that Bm86 consists of approximately 650 amino acids, comprising 6 cysteine residues. Due to the presence of these repeated residues, it is assumed that the protein has a number of epidermal growth-factor (EGF)–like domains that was confirmed with subsequent amino acid sequence analysis (Bastos *et al.*, 2010; Rand *et al.*, 1989). Bm86 contains transmembrane regions during the immature life stage, which are replaced by a GPI-anchor in the adult stages (Rand *et al.*, 1989).

It is hypothesized that Bm86 acts as a cell membrane ligand-gated channel, which allows intracellular communication between neighboring and distant cells mediated by its EGF-like domains and the GPI-anchor (Rand *et al.*, 1989). Vaccination studies and subsequent tick dissection found that antibodies against Bm86 affected the intracellular digestion of blood (Willadsen *et al.*, 1989). These results were supported by studies that showed the presence of undigested red blood cells in Bm86-silenced ticks (Bastos *et al.*, 2010). Functional predictions of this membrane bound antigen also indicated that Bm86 plays a role in cell-cell communication, as well as in blood-digestion (Bastos *et al.*, 2010). In 2007 Nijhof *et al.* set out to investigate Bm86's function, utilizing the reverse genetics approach, RNAi. However, silencing of the Bm86 RNA did not have a significant effect on the fitness of adult *R. microplus* female ticks (Nijhof *et al.*, 2007). In *Haemaphysalis longicornis,* however, posttranscriptional gene silencing of Bm86 resulted in a 20 % reduction of female tick engorgement weight (Liao *et al.*, 2007). Thus, the use of RNAi is not a viable technique to study the function of Bm86.

MRGIALFVAAVSLIVEGTAESSICSDFGNEFCRNAECEVVPGAEDDFVCKCPRDNMYFNAA
EKQCTREYKDTCKECSYGRCVESNPSKASCVCEASDDLTLQCKIKNDYATDCRNRGGTAK
LRTDGFIGATCDCGEWGAMNMTTRNCVPTTCLRPDLTCKDLCEKNLLQRDSRCCQGWNT
ANCSAAPPADSYCSPGSPKGPDGQCINACKTKEAGFVCKHGCRSTKAYECPSGSTVAED
GITCKSISHTVSCTAEQKQTCRPTEDCRVHKGTVLCECPWNQHLVGDTCISDCVDKKCHEE
FMDCGVYMNRQSCYCPWKSRKPGPNVNINECLLNEYYYTVSFTPNISFDSDHCKWYEDRV
LEAIRTSIGKEVFKVEILNCTQDIKARLIAEKPLSKHVLRKLQACEHPIGEWCMMYPKLLIKKN
SATEIEEENLCDSLLKDQEAAYKGQNKCVKVDNLFWFQCADGYTTTYEMTRGRLRRSVCK
AGVSCNENEQSECADKGQIFVYENGKANCQCPPDTKPGEIGCIERTTCNPKEIQECQDKKL
ECVYKNHKAECECPDDHECYREPAKDSCSEEDNGKCQSSGQRCVIENGKAVCKEKSEAT
TAATTTTKAKDKDPGKSSAAAVSATGLLLLLAATSVTAASL

Figure 1.5: Amino acid sequence of Bm86. *Red: signal peptide, blue: epidermal growth (EGF)-like domain and green: glycosylphosphatidylinositol (GPI) anchor sequence (Refer to Figure 1.6).*

Vaccination with recombinant Bm86 was found to protect against tick infestation, but the efficacy of the Bm86 vaccine varied greatly between *R. microplus* strains from 0% to 91% and between different species from 0% to 99.99% (Table 1.3) (De Vos *et al.*, 2001). Immunization with Bm86 led to a reduction in the number of engorging female ticks, their engorgement weight and reproductive capacity by most probably reducing the expression of factors involved in metabolic pathways of feeding, growth and reproduction (De la Fuente *et al.*, 2007b; De la Fuente, Kocan, 2003).

Table 1.3: Effect of vaccination with *R. microplus* recombinant antigens on different tick species.

Tick species and origins	Antigen(strain)	Efficacy (%)	References
Rhipicephalus microplus			
Yeerongpilly	Bm86 (Australian)	89	(Tellam *et al.*, 1992)
Yeerongpilly	Bm86 (Cuban)	75	(Penichet *et al.*, 1994)
Camcord, Cuba	Bm86 (Cuban)	91	(Penichet *et al.*, 1994)
Cenapa	Bm86 (Cuban)	84	(Penichet *et al.*, 1994)
Tuxpan, Mexico	Bm86 (Cuban)	51	(Garcia-Garcia *et al.*, 1999)
Mora	Bm86 (Cuban)	58	(Garcia-Garcia *et al.*, 1999)
Argentina	Bm86 (Cuban)	0	(De la Fuente *et al.*, 2000)
Texas	Bm86 (Cuban)	85,2	(Canales *et al.*, 2009)
Texas	Bm86 (Mozambique)	70	(Canales *et al.*, 2009)
Boophilus annulatus			
Quemado, Mexico	Bm86 (Cuban)	99.98	(De La Fuente *et al.*, 1999)
Rasht-Fouman, Iran	Bm86 (Cuban)	99.99	(De La Fuente *et al.*, 1999)
Texas	Bm86 (Cuban)	99	(Canales *et al.*, 2009)
Texas	Bm86 (Mozambique)	99.6	(Canales *et al.*, 2009)
Boophilus decoloratus			
South Africa	Bm86 (Australian)	70	(De Vos *et al.*, 2001; Odongo *et al.*, 2007)
Hyalomma anatolicum anatolicum			
Iran	Bm86 (Cuban)	97	(De Vos *et al.*, 2001)
Ludhiana, India	Bm86 (Australian)	50	
Hyalomma dromedarii			
Morocco	Bm86 (Australian)	98	
Rhipicephalus appendiculatus			
Harare, Zimbabwe	Bm86 (Australian)	0	
Amblyomma variegatum			
Burkina Faso	Bm86 (Australian)	0	

Variation in efficacy may be due to differences in the amino acid sequence between the recombinant Bm86 vaccine and the native Bm86 expressed in different regions (Table1.3) (Garcia-Garcia *et al.*, 1999; Sossai *et al.*, 2005). It has been found that Bm86 sequences from some regions appear to have a short unique hydrophilic amino acid sequence compared to sequences from other regions. This may cause the vaccine to target other surface exposed epitopes (Freeman *et al.*, 2010). Molecular identification of Bm86 genes in *Hyalomma* species showed that there is more than 86% identity between amino acid sequences in Bm86 homologs from *R. microplus, Hy. marginatum, Hy. excavatum* and *Hy. dromedarii* (Ben Said *et al.*, 2012). The above study was the first to report the presence of two different alleles of Bm86, He86p2a1 and He86p2a2, with 92% identity. This finding allows for the prediction of which epitopes of the *Hyalomma*-Bm86 antigens to use in a vaccine for an improved *Hyalomma* tick infestation control (Ben Said *et al.*, 2012).

Differences in expression levels between various tick life-stages are another possible explanation for the variation in the efficacy of the Bm86 vaccine (Canales *et al.*, 2009; Peconick *et al.*, 2008; Rand *et al.*, 1989). The expression level of Bm86 in *R. microplus* eggs was found to be very low and increases until the tick development reaches the unfed larval stage. Subsequently the expression level drops during blood feeding (Nijhof *et al.*, 2009). Moreover, the conformation of epitopes exposed to the host could have undergone evolutionary changes depending on which host species can be infested (Canales *et al.*, 2009; Nijhof *et al.*, 2010; Peconick *et al.*, 2008; Rand *et al.*, 1989). These changes might result in the production of different anti-Bm86 antibodies (Guerrero *et al.*, 2012b).

Partial protection of the Bm86 vaccine against *Rhipicephalus* and *Hyalomma* tick species suggests that the efficacy of a Bm86-based vaccine may be enhanced when using the orthologous recombinant Bm86 antigen for vaccination (Azhahianambi *et al.*, 2009). It was found that a vaccine based on the Bm86 orthologous gene of *H. a. anatolicum* (Haa86) has an efficacy of 82.3% against *H. a. anatolicum* (Azhahianambi *et al.*, 2009).

Nijhof *et al*. (2010) aimed to identify Bm86 homologs by means of rapid amplification of cDNA ends (RACE) from different tick species. This led to the discovery of a unique protein that is structurally similar to Bm86, based on the presence of multiple

epidermal growth factor (EGF)-like domains, called ATAQ (Bastos *et al.*, 2010; Rand *et al.*, 1989).

1. 1. 9. A novel putative antigen: ATAQ

ATAQ (named after its signature peptide sequence) was identified during the search for Bm86 homologs and was found to be structurally similar to Bm86 (Figure 1.6) (Nijhof *et al.*, 2010). Both proteins have a signal peptide and multiple EGF-like domains; however, ATAQ lacks a GPI-anchor. Additionally, this novel tick protein has a transmembrane region and an intracellular domain (Figure 1.7).

Comparisons between ATAQ and Bm86 did not show a significant sequence similarity, and therefore it is hypothesized that Bm86 and ATAQ have different biological roles in ticks (Nijhof *et al.*, 2010). However, according to *in silico* predictions, antigenic regions of Bm86 and BmATAQ are 44% similar, but the similarity between the amino acid sequences of immunogenic regions is between 9-29%. Hence it is reasonable to assume that anti-Bm86 antibodies would not cross-react with BmATAQ antigens (Nijhof *et al.*, 2010).

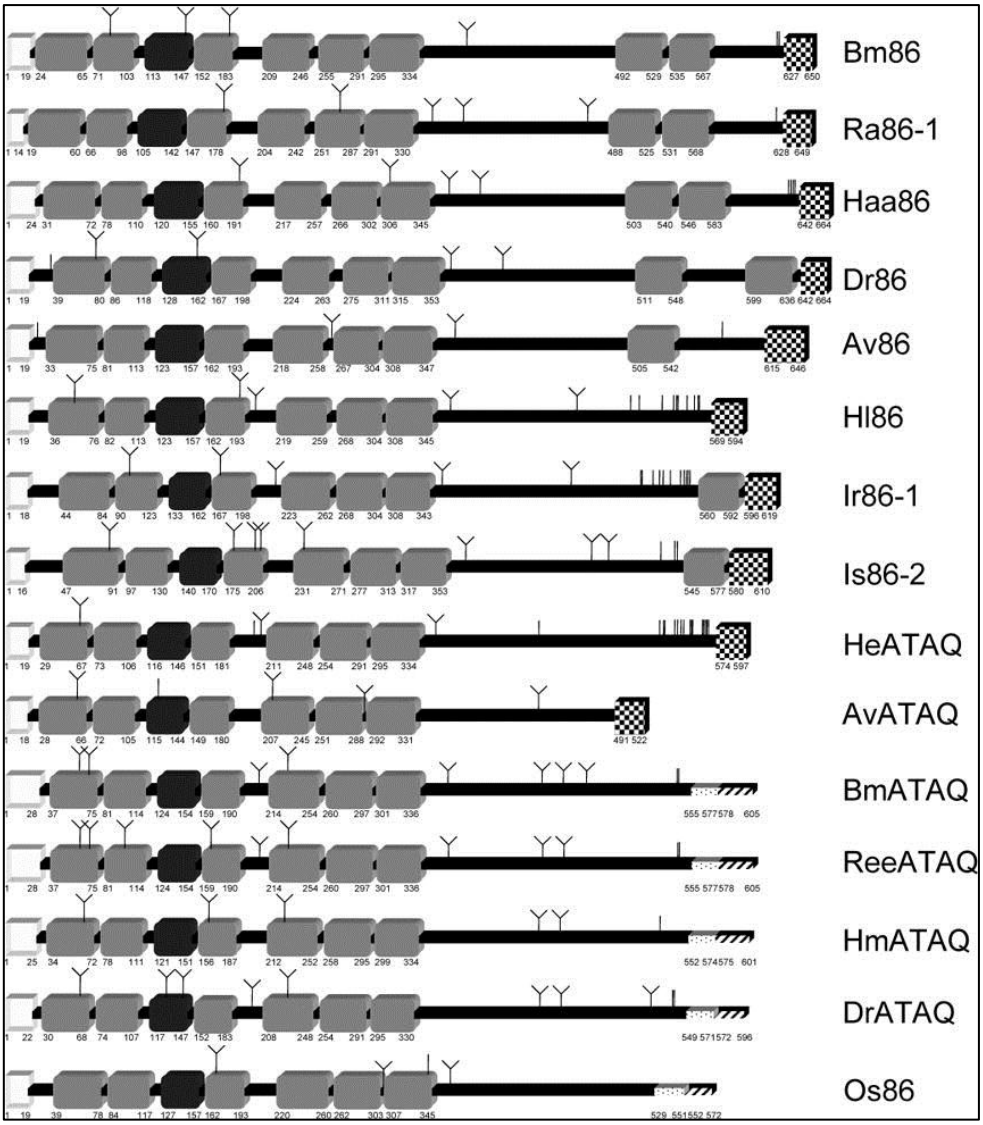


Figure 1.6: Comparison of Bm86 and ATAQ structural characteristics from *R. microplus* with their representative orthologs from other tick species (Nijhof *et al.*, 2010). *White boxes represent the presence of a signal peptide, light and dark gray areas represent EGF-like domains, checker boxes represent a GPI-anchor, dotted highlight represents a transmembrane domain, striped highlight indicates an intracellular domain and Y indicates potential N-linked carbohydrate addition sites. Numbers indicate the respective amino acid residues. Abbreviations correspond to: Bm: Rhipicephalus microplus, Ra: Rhipicephalus annulatus, Haa: Hyalomma anatolicum anatolicum, Dr: Dermacentor reticulatus, Av: Amblyomma variegatum, Hl: Haemaphysalis longicornis, Ir: Ixodes ricinus, Is: Ixodes scapularis, He: Haemaphysalis elliptica, Ree: Rhipicephalus evertsi evertsi, Hm: Hyalomma marginatum, Os: Ornithodoros savignyi.*

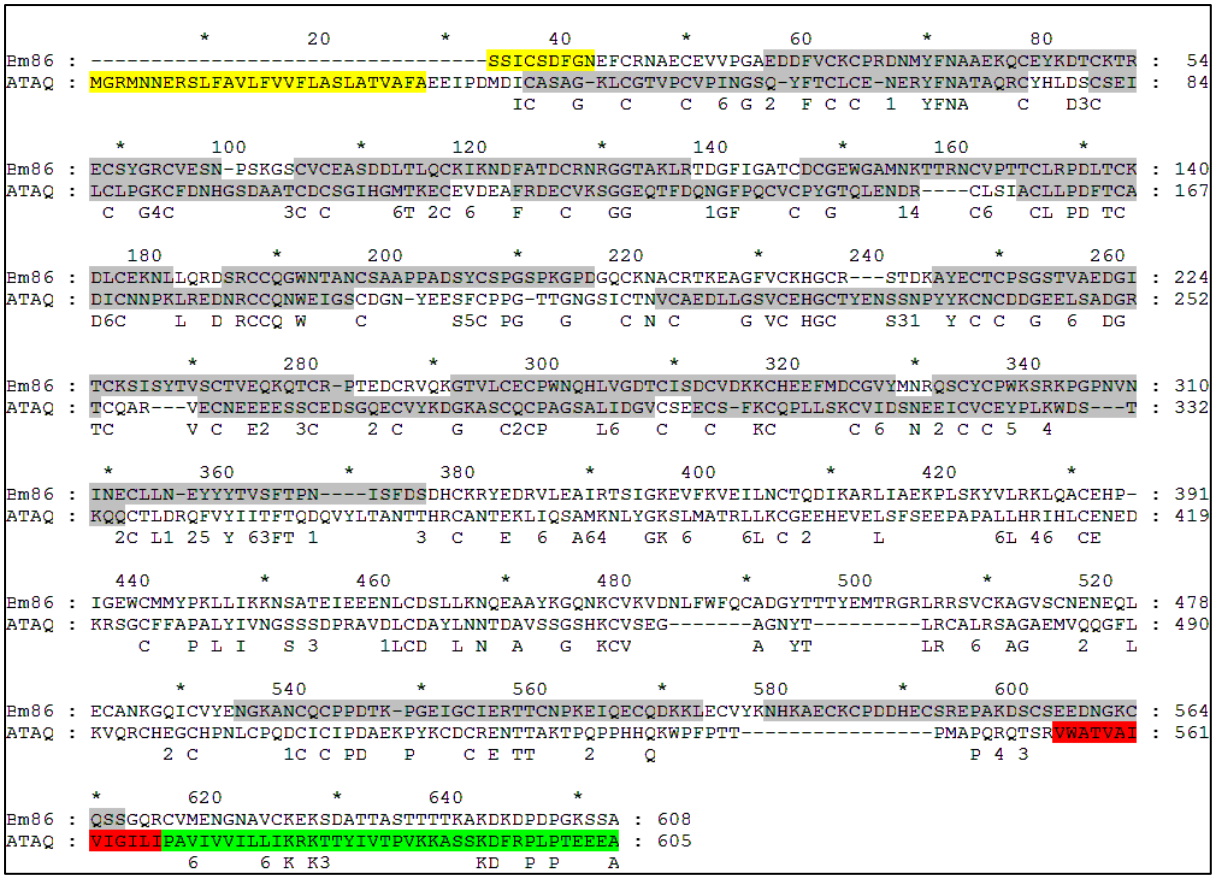


Figure 1.7: Amino acid sequence alignment of *R. microplus* Bm86 and ATAQ. *Signal peptide sequence (yellow), EGF-likedomain (gray), TM: transmembrane sequence (red), and intracellular domain (green) (Motive organization is indicated in Figure 1.6).*

Tissue localization also differs between Bm86 and ATAQ as the Bm86 orthologs are found mostly in midgut cells, while ATAQ is found in the midgut, as well as in the malpighian tubules (Figure 1. 8) (Nijhof *et al.*, 2010). Vaccination with recombinant ATAQ by itself in cattle, however, conferred no protection against *R. microplus* (Dr. Ard Nijhof, personal communication).

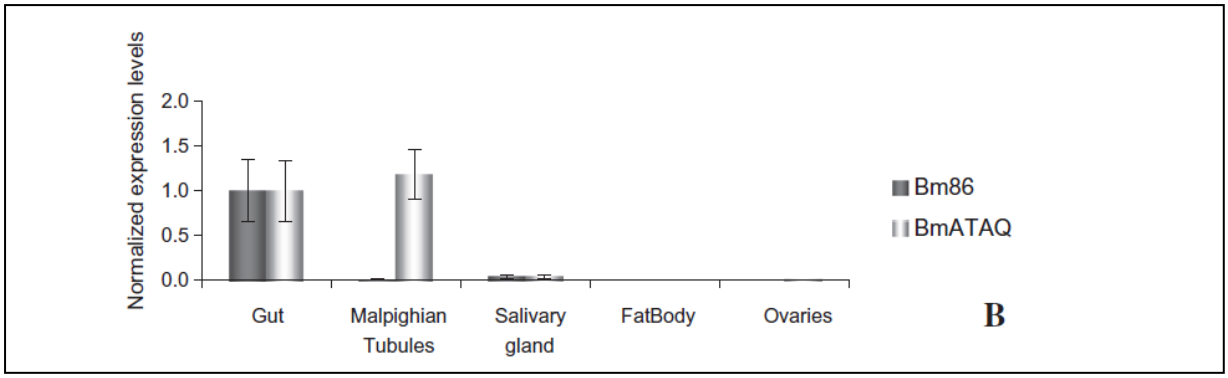


Figure 1.8: Differences between Bm86 and ATAQ expression profiles (Nijhof *et al.*, 2010). *Bm86 is expressed most abundantly in tick gut while ATAQ is found in both gut and malpighian tubules.*

Since Bm86 is the only anti-tick vaccine to ever be commercialized, it is essential to understand its function *in vivo*. It is hypothesized that the vaccine efficacy of Bm86 could be enhanced by targeting proteins interacting with Bm86, or proteins that are involved in the same biological pathway. Therefore, this study set out to analyze protein-protein interactions by means of a yeast two-hybrid screening, using Bm86 and ATAQ from *R. microplus* as bait proteins.

Protein-protein interactions (PPIs) are essential for all biological processes in living organisms. Consequently, to understand the biological role of a protein it is crucial to know the core PPI network. Regions where proteins interact are more conserved than epitopes presented in interspecies' interactomes, as any change in their amino acid region might result in loss of protein interactions vital to the function of large protein complexes. Yeast two-hybrid screening is one of the most reliable methods to generate a PPI dataset using a high-throughput technique. This research therefore focuses on the identification of novel antigens by investigating PPIs of Bm86 and ATAQ with other proteins thereby allowing biological function prediction of the vaccine antigen (Bm86) and novel candidate antigen (ATAQ).

1. 2 HYPOTHESES AND AIMS

Hypotheses

- Bm86 and ATAQ interacting partners can be identified by means of the Yeast two-hybrid system.
- Identification and functional annotation of binding partners will aid the functional annotation of Bm86 and ATAQ.

Aims

1- To obtain a better understanding of the biological roles of Bm86 and ATAQ in *Rhipicephalus microplus* by identifying interacting proteins using a Yeast two-hybrid approach.
2- To use the identified protein binding partner sequences to gain insight into the biological roles of the target proteins using various computational approaches.
3- To confirm the protein-protein interactions by means of immunoprecipitation, western blotting and LC-MS/MS.

1.3 REFERENCES

Almazán C, Moreno-Cantú O, Moreno-Cid JA*, et al.* (2012) Control of tick infestations in cattle vaccinated with bacterial membranes containing surface-exposed tick protective antigens. *Vaccine* **30**, 265-272.

Andreotti R, De León AAP, Dowd SE*, et al.* (2011) Assessment of bacterial diversity in the cattle tick *Rhipicephalus (Boophilus) microplus* through tag-encoded pyrosequencing. *BMC Microbiology* **11**:6.

Atkinson SF, Sarkar S, Avina A, Schuermann JA, Williamson P (2012) Modelling spatial concordance between Rocky Mountain spotted fever disease incidence and habitat probability of its vector *Dermacentor variabilis* (American dog tick). *Geospat Health* **1**, 91-100.

Azhahianambi P, De La Fuente J, Suryanarayana VVS, Ghosh S (2009) Cloning, expression and immunoprotective efficacy of rHaa86, the homologue of the Bm86 tick vaccine antigen, from *Hyalomma anatolicum anatolicum*. *Parasite Immunology* **31**, 111-122.

Barker SC, Murrell A (2002) Phylogeny, evolution and historical zoogeography of ticks: A review of recent progress. *Experimental and Applied Acarology* **28**, 55-68.

Barker SC, Murrell A (2004) Systematics and evolution of ticks with a list of valid genus and species names. *Parasitology* **129**, S15-S36.

Barre N, Uilenberg G (2010) Spread of parasites transported with their hosts: case study of two species of cattle tick. *Revue scientifique et technique (International Office of Epizootics)* **29**, 149-160, 135-147.

Barrero RA, Keeble-Gagnere G, Zhang B*, et al.* (2011) Evolutionary conserved microRNAs are ubiquitously expressed compared to tick-specific miRNAs in the cattle tick *Rhipicephalus (Boophilus) microplus*. *BMC Genomics* **12**:328.

Bastos RG, Ueti MW, Knowles DP, Scoles GA (2010) The *Rhipicephalus (Boophilus) microplus Bm86* gene plays a critical role in the fitness of ticks fed on cattle during acute *Babesia bovis* infection. *Parasites and Vectors* **3**:111.

Ben Said M, Galai Y, Mhadhbi M*, et al.* (2012) Molecular characterization of Bm86 gene orthologs from Hyalomma excavatum, Hyalomma dromedarii and Hyalomma marginatum marginatum and comparison with a vaccine candidate from Hyalomma scupense. *Veterinary Parasitology* **190**(1-2), 230-40.

Beugnet F, Franc M (2012) Insecticide and acaricide molecules and/or combinations to prevent pet infestation by ectoparasites. *Trends in Parasitology* **28**, 267-279.

Bock R, Jackson L, De Vos A, Jorgensen W (2004) Babesiosis of cattle. *Parasitology* **129**, S247-S269.

Boretti FS, Perreten A, Meli ML*, et al.* (2009) Molecular investigations of *Rickettsia helvetica* infection in dogs, foxes, humans, and Ixodes ticks. *Applied and Environmental Microbiology* **75**, 3230-3237.

Botelho-Nevers E, Socolovschi C, Raoult D, Parola P (2012) Treatment of Rickettsia spp. infections: A review. *Expert Review of Anti-Infective Therapy* **10**, 1425-1437.

Botha M (2013) *Two-hybrid analysis and attempted expression of elongation factor 1α from the cattle tick, Rhipicephalus microplus*, University of Pretoria.

Brake DK, Perez De Leon AA (2012) Immunoregulation of bovine macrophages by factors in the salivary glands of *Rhipicephalus microplus*. *Parasites and Vectors* **5**:38.

Bratton RL, Corey GR (2005) Tick-borne disease. *American Family Physician* **71**, 2323-2332.

Brossard M, Wikel SK (2004) Tick immunobiology. *Parasitology* **129**, S161-S176.

Brown WC (2001) Molecular approaches to elucidating innate and acquired immune responses to *Babesia bovis*, a protozoan parasite that causes persistent infection. *Veterinary Parasitology* **101**, 233-248.

Burg JG (2001) Seasonal activity and spatial distribution of host-seeking adults of the tick Dermacentor variabilis. *Medical and Veterinary Entomology* **15**, 413-421.

Camargo Mathias MI, Scopinho Furquim KC, Nues PH (2011) Immunomodulatory effects of tick saliva. *Information Systems Journal* **8**, 231-240.

Canales M, Almazan C, Naranjo V, Jongejan F, de la Fuente J (2009) Vaccination with recombinant *Boophilus annulatus* Bm86 ortholog protein, Ba86, protects cattle against *B. annulatus* and *B. microplus* infestations. *BMC Biotechnology* **9**:29.

De Castro JJ (1997) Sustainable tick and tickborne disease control in livestock improvement in developing countries. *Veterinary Parasitology* **71**, 77-97.

De la Fuente J, Almazan C, Canales M*, et al.* (2007) A ten-year review of commercial vaccine performance for control of tick infestations on cattle. *Animal Health Research Reviews* **8**, 23 - 28.

De la Fuente J, Kocan KM (2003) Advances in the identification and characterization of protective antigens for recombinant vaccines against tick infestations. *Expert Review of Vaccines* **2**, 583-593.

De la Fuente J, Kocan M (2006) Strategies for development of vaccines for control of ixodid tick species. *Parasite Immunology* **28**, 275-283.

De La Fuente J, Rodriguez M, Montero C*, et al.* (1999) Vaccination against ticks (Boophilus spp.): The experience with the Bm86-based vaccine Gavac((TM)). *Genetic Analysis - Biomolecular Engineering* **15**, 143-148.

De la Fuente JD, Rodriguez M, Garcia-Garcia JC (2000) Immunological control of ticks through vaccination with *Boophilus microplus* gut antigens. In: *Annals of the New York Academy of Sciences* **916**, pp. 617-621.

De Vos S, Zeinstra L, Taoufik A, Willadsen P, Jongejan F (2001) Evidence for the utility of the Bm86 antigen from *Boophilus microplus* in vaccination against other tick species. *Experimental and Applied Acarology* **25**, 245-261.

Decrem Y, Mariller M, Lahaye K*, et al.* (2008) The impact of gene knock-down and vaccination against salivary metalloproteases on blood feeding and egg laying by *Ixodes ricinus*. *International Journal for Parasitology* **38**, 549-560.

Dobson SJ, Barker SC (1999) Phylogeny of the Hard Ticks (Ixodidae) Inferred from 18S rRNA Indicates That the Genus Aponommals Paraphyletic. *Molecular Phylogenetics and Evolution* **11**, 288-295.

Domingos A, Antunes S, Borges L, Rosario VEd (2013) Approaches towards tick and tick-borne diseases control. *Revista da Sociedade Brasileira de Medicina Tropical*, 00-00.

Donohue KV, Khalil SMS, Sonenshine DE, Roe RM (2009) Heme-binding storage proteins in the Chelicerata. *Journal of Insect Physiology* **55**, 287-296.

Embers ME, Narasimhan S (2013) Vaccination against Lyme disease: past, present, and future. *Frontiers in Cellular and Infection Microbiology* **3**:6.

Ergonul O (2012) Crimean-Congo hemorrhagic fever virus: new outbreaks, new discoveries. *Current Opinion in Virology* **2**, 215-220.

Estrada-Pena A (2008) Climate, niche, ticks, and models: What they are and how we should interpret them. *Parasitology Research* **103**, S87-S95.

Estrada-Peña A, Estrada-Sánchez A, Estrada-Sánchez D (2012a) Occurrence patterns of afrotropical ticks (Acari: Ixodidae) in the climate space are not correlated with their taxonomic relationships. *PLoS ONE* **7**(5).

Estrada-Peña A, García Z, Sánchez HF (2006) The distribution and ecological preferences of *Boophilus microplus* (Acari: Ixodidae) in Mexico. *Experimental and Applied Acarology* **38**, 307-316.

Estrada-Peña A, Venzal JM, Nava S*, et al.* (2012b) Reinstatement of *Rhipicephalus (Boophilus) australis* (Acari: Ixodidae) with redescription of the adult and larval stages. *Journal of Medical Entomology* **49**, 794-802.

Fernandes ÃKK, Bittencourt VREP (2008) Entomopathogenic fungi against South American tick species. *Experimental and Applied Acarology* **46**, 71-93.

Fernandes EKK, Bittencourt VREP, Roberts DW (2012) Perspectives on the potential of entomopathogenic fungi in biological control of ticks. *Experimental Parasitology* **130**, 300-305.

Francischetti IMB, Mather TN, Ribeiro JMC (2003) Cloning of a salivary gland metalloprotease and characterization of gelatinase and fibrin(ogen)lytic activities in the saliva of the Lyme disease tick vector *Ixodes scapularis*. *Biochemical and Biophysical Research Communications* **305**, 869-875.

Freeman JM, Davey RB, Kappmeyer LS, Kammlah DM, Olafson PU (2010) Bm86 midgut protein sequence variation in South Texas cattle fever ticks. *Parasites and Vectors* **3**:101.

Gage KL, Gilmore RD, Karstens RH, Schwan TG (1992) Detection of *Rickettsia rickettsii* in saliva, hemolymph and triturated tissues of infected *Dermacentor andersoni* ticks by polymerase chain reaction. *Molecular and Cellular Probes* **6**, 333-341.

Garcia-Garcia JC, Gonzalez IL, Gonzalez DM*, et al.* (1999) Sequence variations in the *Boophilus microplus* Bm86 locus and implications for immunoprotection in cattle vaccinated with this antigen. *Experimental and Applied Acarology* **23**, 883-895.

Gomez-Diaz E, Morris-Pocock JA, Gonzalez-Solis J, McCoy KD (2012) Trans-oceanic host dispersal explains high seabird tick diversity on Cape Verde islands. *Biology Letters* **8**, 616-619.

Gough JM, Kemp DH (1993) Localization of a low abundance membrane protein (Bm86) on the gut cells of the cattle tick *Boophilus microplus* by immunogold labeling. *Journal of Parasitology* **79**, 900-907.

Gourlay LJ, Colombo G, Soriani M*, et al.* (2009) Why is a protective antigen protective? *Human Vaccines* **5**(12): 872-5.

Graca-Souza AV, Maya-Monteiro C, Paiva-Silva GO*, et al.* (2006) Adaptations against heme toxicity in blood-feeding arthropods. *Insect Biochemistry and Molecular Biology* **36**, 322-335.

Graham J, Stockley K, Goldman RD (2011) Tick-borne illnesses: A CME update. *Pediatric Emergency Care* **27**, 141-147.

Gray J, Zintl A, Hildebrandt A, Hunfeld KP, Weiss L (2010) Zoonotic babesiosis: Overview of the disease and novel aspects of pathogen identity. *Ticks and Tick-borne Diseases* **1**, 3-10.

Guerrero FD, Miller RJ, Perez de Leon AA (2012a) Cattle tick vaccines: Many candidate antigens, but will a commercially viable product emerge? *International Journal for Parasitology* **42**, 421-427.

Guerrero FD, Miller RJ, Pérez de León AA (2012b) Cattle tick vaccines: Many candidate antigens, but will a commercially viable product emerge? *International Journal for Parasitology* **42**, 421-427.

Harwood D, Aoki H, Lee Y-D, Fellers JF, White JL (1979) Solution and Rheological Properties of Poly (m-Phenyleneisophthalamide) in Dimethlacetamide/LiCl. *Journal of Applied Polymer Science* **23**, 2155-2168.

Hasle G, Bjune GA, Christensson D*, et al.* (2010) Detection of *Babesia divergens* in southern Norway by using an immunofluorescence antibody test in cow sera. *Acta Veterinaria Scandinavica* **52**:55.

Hill CA, Wikel SK (2005) The *Ixodes scapularis* Genome Project: An opportunity for advancing tick research. *Trends in Parasitology* **21**, 151-153.

Homer MJ, Aguilar-Delfin I, Telford Iii SR, Krause PJ, Persing DH (2000) Babesiosis. *Clinical Microbiology Reviews* **13**, 451-469.

Hovius JWR, Levi M, Fikrig E (2008) Salivating for Knowledge: Potential Pharmacological Agents in Tick Saliva. *PLoS Medicine* **5**, e43.

Hovius JWR, Ramamoorthi N, Van't Veer C*, et al.* (2007) Identification of Salp15 homologues in *Ixodes ricinus* ticks. *Vector-Borne and Zoonotic Diseases* **7**, 296-303.

Jensen K, de Miranda Santos IKF, Glass EJ (2007) Using genomic approaches to unravel livestock (host)-tick-pathogen interactions. *Trends in Parasitology* **23**, 439-444.

Jongejan F, Nene V, de la Fuente J, Pain A, Willadsen P (2007) Advances in the genomics of ticks and tick-borne pathogens. *Trends in Parasitology* **23**, 391-396.

Jongejan F, Uilenberg G (2004) The global importance of ticks. *Parasitology* **129**, S3-S14.

Kalinova Z, Cislakova L, Halanova M (2009) Ehrlichiosis/Anaplasmosis. *Klinicka Mikrobiologie a Infekcni Lekarstv* **15**, 210-213.

Kalume MK, Losson B, Angenot L*, et al.* (2012) Rotenoid content and in vitro acaricidal activity of *Tephrosia vogelii* leaf extract on the tick *Rhipicephalus appendiculatus*. *Veterinary Parasitology* **190**, 204-209.

Kamau L, Skilton RA, Odongo DO*, et al.* (2011) Differential transcription of two highly divergent gut-expressed Bm86 antigen gene homologues in the tick *Rhipicephalus appendiculatus* (Acari: Ixodida). *Insect Molecular Biology* **20**, 105-114.

Kazimirova M (2008) Pharmacologically active compounds in tick salivary glands. *Universal Decimal Classification* **12**, 281-296.

Kelly PJ, Xu C, Lucas H*, et al.* (2013) Ehrlichiosis, Babesiosis, Anaplasmosis and Hepatozoonosis in Dogs from St. Kitts, West Indies. *PLoS ONE* **8** (1):e53450.

Klasco R (2002) Colorado tick fever. *Medical Clinics of North America* **86**, 435-440.

Klompen JSH, Black Iv WC, Keirans JE, Oliver Jr JH (1996) Evolution of ticks, pp. 141-161.

Kocan KM, De la Fuente J, Guglielmone AA, Melendez RD (2003) Antigens and Alternatives for Control of *Anaplasma marginale* Infection in Cattle. *Clinical Microbiology Reviews* **16**, 698-712.

Kuno G, Chang GJJ (2005) Biological transmission of arboviruses: Reexamination of and new insights into components, mechanisms, and unique traits as well as their evolutionary trends. *Clinical Microbiology Reviews* **18**, 608-637.

Lawrence E (2005) *Henderson's Dictionary of Biology*, Thirteenth edn. Pearson Prentice Hall, England

Leblebicioglu H, Bodur H, Dokuzoguz B*, et al.* (2012) Case management and supportive treatment for patients with Crimean-Congo hemorrhagic fever. *Vector-Borne and Zoonotic Diseases* **12**, 805-811.

Li AY, Chen AC, Miller RJ, Davey RB, George JE (2007) Acaricide resistance and synergism between permethrin and amitraz against susceptible and resistant strains of *Boophilus microplus* (Acari: Ixodidae). *Pest Management Science* **63**, 882-889.

Liao M, Zhou J, Hatta T*, et al.* (2007) Molecular characterization of *Rhipicephalus (Boophilus) microplus* Bm86 homologue from *Haemaphysalis longicornis* ticks. *Veterinary Parasitology* **146**, 148-157.

Lindquist L, Vapalahti O (2008) Tick-borne encephalitis. *The Lancet* **371**, 1861-1871.

Lynen G, Zeman P, Bakuname C*, et al.* (2008) Shifts in the distributional ranges of Boophilus ticks in Tanzania: Evidence that a parapatric boundary between *Boophilus microplus* and *B. decoloratus* follows climate gradients. *Experimental and Applied Acarology* **44**, 147-164.

Madder M, Thys E, Achi L, Toure A, De Deken R (2011) *Rhipicephalus (Boophilus) microplus*: A most successful invasive tick species in West-Africa. *Experimental and Applied Acarology* **53**, 139-145.

Madder M, Thys E, Geysen D, Baudoux C, Horak I (2007) *Boophilus microplus* ticks found in West Africa. *Experimental and Applied Acarology* **43**, 233-234.

Mans BJ, de Klerk D, Pienaar R, de Castro MH, Latif AA (2012) The Mitochondrial Genomes of *Nuttalliella namaqua* (Ixodoidea: Nuttalliellidae) and *Argas africolumbae* (Ixodoidae: Argasidae): Estimation of Divergence Dates for the Major Tick Lineages and Reconstruction of Ancestral Blood-Feeding Characters. *PLoS ONE* **7**(11): e49461.

Mans BJ, de Klerk D, Pienaar R, Latif AA (2011) *Nuttalliella namaqua*: A living fossil and closest relative to the ancestral tick lineage: Implications for the evolution of blood-feeding in ticks. *PLoS ONE* **6**(8): e23675.

Mans BJ, Gothe R, Neitz AWH (2004) Biochemical perspectives on paralysis and other forms of toxicoses caused by ticks. *Parasitology* **129**, S95-S111.

Mans BJ, Louw AI, Neitz AWH (2002) Evolution of hematophagy in ticks: Common-origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus *Ornithodoros. Molecular Biology and Evolution* **19**, 1695-1705.

Minniear TD, Buckingham SC (2009) Managing Rocky Mountain spotted fever. *Expert Review of Anti-Infective Therapy* **7**, 1131-1137.

Nadelman RB, Wormser GP (2013) Reinfection versus Relapse in Lyme Disease. *New England Journal of Medicine* **368**, 1063-1064.

Nieto NC, Teglas MB, Stewart KM, Wasley T, Wolff PL Detection of relapsing fever spirochetes (*Borrelia hermsii* and *Borrelia coriaceae*) in free-ranging mule deer (*Odocoileus hemionus*) from Nevada, United States. *Vector-Borne and Zoonotic Diseases* **12**, 99-105.

Nijhof AM, Balk JA, Postigo M, Jongejan F (2009) Selection of reference genes for quantitative RT-PCR studies in *Rhipicephalus (Boophilus) microplus* and Rhipicephalus appendiculatus ticks and determination of the expression profile of Bm86. *BMC Molecular Biology* **10**.

Nijhof AM, Balk JA, Postigo M*, et al.* (2010) Bm86 homologues and novel ATAQ proteins with multiple epidermal growth factor (EGF)-like domains from hard and soft ticks. *International Journal for Parasitology* **40**, 1587-1597.

Nijhof AM, Taoufik A, de la Fuente J*, et al.* (2007) Gene silencing of the tick protective antigens, Bm86, Bm91 and subolesin, in the one-host tick *Boophilus microplus* by RNA interference. *International Journal for Parasitology* **37**, 653-662.

Nunn CL, Ezenwa VO, Arnold C, Koenig WD (2011) Mutualism or parasitism? using a phylogenetic approach to characterize the oxpecker-ungulate relationship. *Evolution* **65**, 1297-1304.

Nuttall PA, Labuda M (2004) Tick-host interactions: Saliva-activated transmission. *Parasitology* **129**, S177-S189.

Nuttall PA, Trimnell AR, Kazimirova M, Labuda M (2006) Exposed and concealed antigens as vaccine targets for controlling ticks and tick-borne diseases. *Parasite Immunology* **28**, 155-163.

Nyangiwe N, Goni S, Herve-Claude LP, Ruddat I, Horak IG (2011) Ticks on pastures and on two breeds of cattle in the Eastern Cape province, South Africa. *Ondestepoort Journal of Veterinary Research.* **78**, 1-9.

Odongo D, Kamau L, Skilton R*, et al.* (2007) Vaccination of cattle with TickGARD induces cross-reactive antibodies binding to conserved linear peptides of Bm86 homologues in *Boophilus decoloratus*. *Vaccine* **25**, 1287-1296.

Oliver JH (1989) Biology and systemics of ticks (Acari:Ixodida). *Annual Review of Ecology and Systematics* **20**, 397-430.

Olwoch JM, Reyers B, Engelbrecht FA, Erasmus BFN (2008) Climate change and the tick-borne disease, Theileriosis (East Coast fever) in sub-Saharan Africa. *Journal of Arid Environments* **72**, 108-120.

Pagel Van Zee J, Geraci NS, Guerrero FD*, et al.* (2007) Tick genomics: The Ixodes genome project and beyond. *International Journal for Parasitology* **37**, 1297-1305.

Peconick AP, Sossai S, GirÃo FA*, et al.* (2008) Synthetic vaccine (SBm7462) against the cattle tick *Rhipicephalus (Boophilus) microplus*: Preservation of immunogenic determinants in different strains from South America. *Experimental Parasitology* **119**, 37-43.

Penichet M, Rodriguez M, Castellano O*, et al.* (1994) Detection of Bm86 antigen in different strains of *Boophilus microplus* and effectiveness of immunization with recombinant Bm86. *Parasite Immunology* **16**, 493-500.

Perez de Leon A, Strickman D, Knowles D*, et al.* (2010) One Health approach to identify research needs in bovine and human babesioses: workshop report. *Parasites & Vectors* **3**, 36.

Piesman J, Eisen L (2008) Prevention of tick-borne diseases. In: *Annual Review of Entomology*, pp. 323-343.

Rand KN, Moore T, Sriskantha A*, et al.* (1989) Cloning and expression of a protective antigen from the cattle tick *Boophilus microplus*. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 9657-9661.

Reichard MV, Edwards AC, Meinkoth JH*, et al.* (2010) Confirmation of *Amblyomma americanum* (Acari: Ixodidae) as a vector for *Cytauxzoon felis* (Piroplasmorida: Theileriidae) to domestic cats. *Journal of Medical Entomology* **47**, 890-896.

Ribeiro JMC (1995) Blood-feeding arthropods: Live syringes or invertebrate pharmacologists? *Infectious Agents and Disease* **4**, 143-152.

Ribeiro JMC, Alarcon-Chaidez F, B. Francischetti IM*, et al.* (2005) An annotated catalog of salivary gland transcripts from *Ixodes scapularis* ticks. *Insect Biochemistry and Molecular Biology* **36**, 111-129.

Sauer JR, Essenberg RC, Bowman AS (2000) Salivary glands in ixodid ticks: Control and mechanism of secretion. *Journal of Insect Physiology* **46**, 1069-1078.

Simons SM, Junior PLDS, Faria F*, et al.* (2011) The action of Amblyomma cajennense tick saliva in compounds of the hemostatic system and cytotoxicity in tumor cell lines. *Biomedicine and Pharmacotherapy* **65**, 443-450.

Sonenshine DE (1991) *Biology of ticks.*Oxford University Press, New York, Oxford.

Sossai S, Peconick AP, Sales Jr PA*, et al.* (2005) Polymorphism of the bm86 gene in South American strains of the cattle tick *Boophilus microplus*. *Experimental and Applied Acarology* **37**, 199-214.

Sprong H, Wielinga PR, Fonville M*, et al.* (2009) Ixodes ricinus ticks are reservoir hosts for *Rickettsia helvetica* and potentially carry flea-borne Rickettsia species. *Parasites and Vectors* **2**:41.

Stephen C, Artsob H, Bowie WR*, et al.* (2004) Perspectives on emerging zoonotic disease research and capacity building in Canada. *Canadian Journal of Infectious Diseases* **15**, 339-344.

Suarez CE, Noh S (2011) Emerging perspectives in the research of bovine babesiosis and anaplasmosis. *Veterinary Parasitology* **180**, 109-125.

Teglas MB, Mapes S, Hodzic E, Nieto NC (2011) Co-infection of *Ornithodoros coriaceus* with the relapsing fever spirochete, *Borrelia coriaceae*, and the agent of epizootic bovine abortion. *Medical and Veterinary Entomology* **25**, 337-343.

Tellam RL, Smith D, Kemp DH, Willadsen P (1992) Vaccination against ticks.*Animal Parasite Control Utilizing Biothechnology*, ed. 303-331.

Terkawi MA, Thekisoe OMM, Katsande C*, et al.* (2011) Serological survey of *Babesia bovis* and *Babesia bigemina* in cattle in South Africa. *Veterinary Parasitology* **182**, 337-342.

Tonnesen MH, Penzhorn BL, Bryson NR, Stoltsz WH, Masibigiri T (2004) Displacement of *Boophilus decoloratus* by *Boophilus microplus* in the Soutpansberg region, Limpopo Province, South Africa. *Experimental and Applied Acarology* **32**, 199-208.

Troughton DR, Levin ML (2007) Life cycles of seven ixodid tick species (Acari: Ixodidae) under standardized laboratory conditions. *Journal of Medical Entomology* **44**, 732-740.

Ullmann AJ, Lima CMR, Guerrero FD, Piesman J, Black Iv WC (2005) Genome size and organization in the blacklegged tick, *Ixodes scapularis* and the Southern cattle tick, *Boophilus microplus*. *Insect Molecular Biology* **14**, 217-222.

Valenzuela JG (2004) Exploring tick saliva: From biochemistry to 'sialomes' and functional genomics. *Parasitology* **129**, S83-S94.

Vial L (2009) Biological and ecological characteristics of soft ticks (Ixodida: Argasidae) and their impact for predicting tick and associated disease distriblftion. *Parasite* **16**, 191-202.

Walker AR, Camicas J-L, Estrada-Pena A*, et al.* (2003) Ticks of Domestik Animals in Africa – a guide to Identification of species. Bioscience reports, Edinburgh.

Wang M, Guerrero FD, Pertea G, Nene VM (2007) Global comparative analysis of ESTs from the southern cattle tick, *Rhipicephalus (Boophilus) microplus*. *BMC Genomics* **8**.

Wikswo ME, Hu R, Dasch GA*, et al.* (2008) Detection and identification of spotted fever group rickettsiae in dermacentor species from Southern California. *Journal of Medical Entomology* **45**, 509-516.

Willadsen P (2004) Anti-tick vaccines. *Parasitology* **129**, S367-S387.

Willadsen P (2006a) Tick control: Thoughts on a research agenda. *Veterinary Parasitology* **138**, 161-168.

Willadsen P (2006b) Vaccination against ectoparasites. *Parasitology* **133**, S9-S25.

Willadsen P, Riding GA, McKenna RV*, et al.* (1989) Immunologic control of a parasitic arthropod. Identification of a protective antigen from *Boophilus microplus*. *Journal of Immunology* **143**, 1346-1351.

Zintl A, Mulcahy G, Skerrett HE, Taylor SM, Gray JS (2003) *Babesia divergens*, a Bovine Blood Parasite of Veterinary and Zoonotic Importance. *Clinical Microbiology Reviews* **16**, 622-636.

Zivkovic Z, Esteves E, Almazan C*, et al.* (2010) Differential expression of genes in salivary glands of male *Rhipicephalus (Boophilus) microplus* in response to infection with *Anaplasma marginale*. *BMC Genomics* **11**:186.

CHAPTER 2

Identification and validation of interacting partners of Bm86 and BmATAQ from *R. microplus* using the Yeast two-hybrid system

2.1. INTRODUCTION

Without protein-protein interactions, maintenance of virtually all cellular mechanisms in the cell would not be possible. Protein-protein interactions can be extra- or intracellular and include cell-cell adhesion, ligand-receptor interactions, virus-host recognition and immune responses such as antibody-antigen recognition (Kim *et al.*, 2011; Vinayagam *et al.*, 2011; Young, 1998). Intracellular interactions include the formation of multiprotein complexes leading to the formation of cytoskeletal elements, transcription factor assembly and signal transduction events (Kim *et al.*, 2011). Protein-protein interactions therefore form the basis of most biological processes.

Proteins interact because of shape, charge, and/ or complementarity in polarity (Jones, 2012). Hydrophobic pockets can be formed by hydrophobic amino acids such as phenylalanine (Phe) and tryptophan (Trp) into which a second hydrophobic area can dock. Positive amino acids interact with negative amino acids, such asparginine (Arg) and lysine (Lys) residues that may interact with an aspartate (Asp) residue. Hydration further stabilizes the interactions with up to 50 molecules of water being involved (Alberts *et al*., 2002). Based on the latter, it is evident that the chemical nature, the type of solvent, its ionic strength and pH affects the extent to which a protein can interact (Jones, 2012).

2.1.1. Studying protein-protein interactions

Classical theory suggested that one protein has one function. However, with the rise of systems biology, a protein's biological role is now summarized as a combination of all the proteins it interacts with (Cox, Mann, 2011). Complete interaction mapping of any novel protein requires information on the interacting partners (Wilson, 2004).

Therefore, identifying interacting partners with known function may thus help to infer an unknown protein's function (Sleator, 2012). There are several approaches to predict protein function such as guilt-by-association (GBA, which focuses on predicting protein interactions), regulatory associations or the effect of mutations (Gillis, Pavlidis, 2012).

Various *in vitro* techniques for the identification of protein-protein interactions can be used. These include chemical cross-linking, affinity chromatography, immunoprecipitation or fluorescence gel retardation (Table 2.1). The primary disadvantage of these techniques is that they require large amounts of pure recombinant protein and in most cases only a single interaction is analyzed at any given time (Miernyk, Thelen, 2008). Optimal detection of transient protein interactions is another limitation, especially if stringent procedures are used. Subsequently weak protein interactions cannot be identified for a given protein (Phizicky, Fields, 1995).

Table 2.1: Methods to detect protein–protein interactions (adapted from Howell *et al*., 2006).

Protein Interaction Method	Description
2-Hybrid-based Approaches	It is an *in vivo* system: usually requiring the use of library fusions (high throughput), where the bait protein of interest is fused to a DNA-binding domain and prey proteins of interest are attached to an activation domain. Selection of co-expression of interacting bait–prey fusions can be preceded according to a phenotypic effect.
Coimmunoprecipitation (co-IP)	Use of bait-specific antibody to co-precipitate the bait protein along with any bait-associated proteins.
Affinity Chromatography	Immobilization of various fusions such as GST, TAP tags, His6 or otherwise a covalently immobilized bait protein; requires the use of solid support matrices which can interact with the fusion proteins of interest. Targeted prey proteins can be applied to the bait-bound matrix, and the complex can be eluted across disruption of the matrix–tag interaction.
Phage display	It is a library-based system; involves generation of prey peptides fused to viral coat proteins; isolation of fusions interacting with the bait protein of interest will give rise to plaques, from which the amplified fusion protein can be isolated and subjected to further screening.
Cross-linking	Some of them are cleavable, to separate protein complexes for analysis. Allows detection of weaker interactions. It is possible for applications both *in vivo* and *in vitro*.
Protein chip arrays	Requires immobilization of bait proteins of interest onto a chip surface, followed by exposure to potential prey proteins and detection of interactions through MS, or colorimetric detection.
Far-western blotting/ELISA	A prey proteins of interest is immobilized to a surface, which is supported by incubation with bait protein of interest (primary antibody), and detection of bait protein binding typically through an antibody to the bait protein/ fusion tag (secondary antibody).
Biophysical Techniques	(a) Surface plasmon resonance (b) Fluorescence resonance energy transfer

In addition to the above mentioned *in vitro* methods, high-throughput library-based, as well as *in vivo* techniques can also be used to study protein-protein interactions. These include the phage display and the yeast two-hybrid systems (Table 2.1) (McCafferty *et al.*, 1990; Smith, 1985; Young, Davis, 1983). Library screens are generally performed in yeast, as their short generation time results in a higher throughput compared to most other methods. In addition, yeast provides co-factors and is capable of performing posttranslational modifications that may be crucial for the functional interaction between proteins (Carey *et al.*, 2012). In this study, the yeast two-hybrid screening system was used to identify interacting partners of Bm86 and ATAQ.

2.1.2. Yeast two-hybrid system

The yeast two-hybrid system was developed in 1987 by Stanley Field and published in 1989 (Fields, Song, 1989). The yeast proteins transcriptional activator GAL4, protein kinases SNF1 and SNF4 were studied based on the expression of genes encoding galactose utilizing enzymes. GAL4 is a positive regulator of galactose-induced gene expression. SNF1 is a regulator of the shift taking place from fermentation to aerobic metabolism and interaction with the SNF4 protein is involved in cell cycle regulation in yeast (Hedbacker, Carlson, 2008; Schuller, 2003). Results showed that transcriptional activation needed two functionally critical and separable domains, an N-terminal DNA-binding domain that binds to certain sequences on the genomic DNA and a C-terminal activation domain to activate transcription (Miller, Stagljar, 2004).

The yeast two-hybrid system is therefore based on the interaction of the GAL4 activation domain (AD) and the DNA binding domain (BD) in yeast. These functional domains are cloned into different vectors into which genes of interest are cloned so that the encoding gene products will express a fused construct of the interacting protein with either the BD or the AD. Therefore, the end products of the two plasmids will interact and activate reporter gene function (Figure 2.1). The gene product of the activated reporter gene allows cells in which interaction occurred to finally grow on selective media lacking essential nutrients. In the case of lacking interaction, the reporter gene will not be expressed, therefore no product will be produced and thus no cell growth will be possible (Miller, Stagljar, 2004).

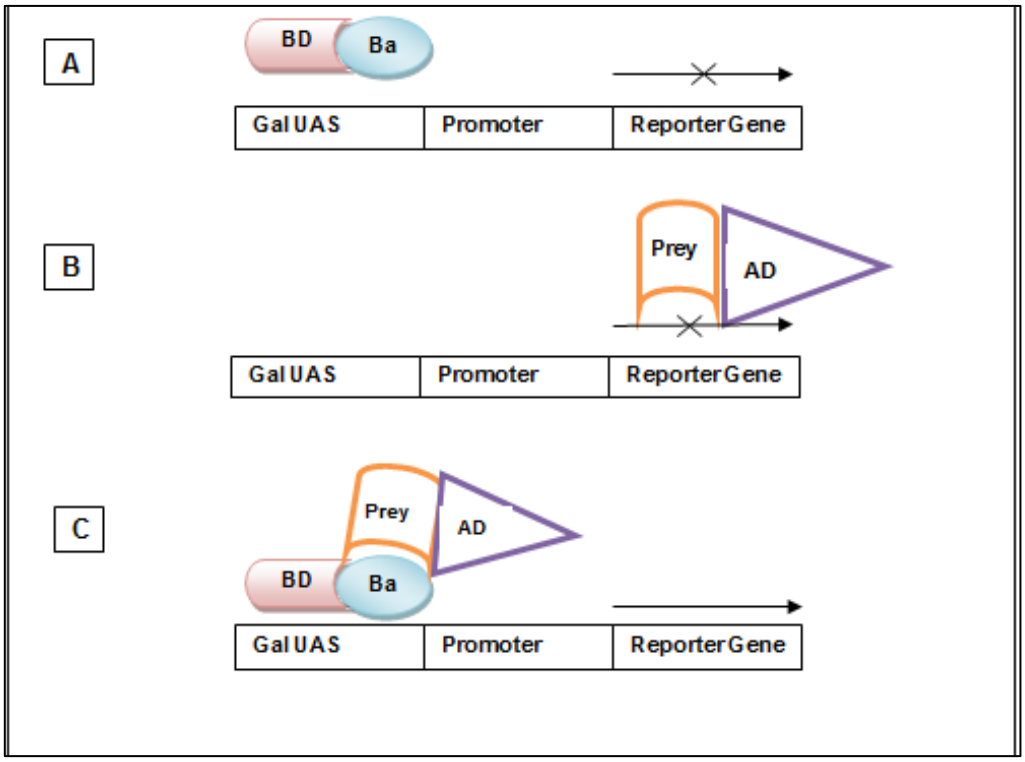


Figure 2.1: Schematic diagram of the Gal-based two-hybrid system. *(**A**) Activation of the reporter gene cannot occur with only DNA-BD/bait hybrid. (**B**) The AD/library fusion protein cannot bind to Gal UAS and activate transcription of the reporter gene without DNA-BD/bait hybrid. (**C**) Transcription of the reporter gene will only be possible with interaction between the bait and library proteins in vivo.*

Yeast strains

Numerous yeast strains are available for use in two-hybrid screening with all of them based on the control of the GAL4 promoter. In this study, the AH109 and Gold yeast strains were used. Both yeast strains enable the elimination of false positives due to the presence of three reporter genes (*HIS3*, *ADE2* and *lacZ*) that increase stringency. The yeast strain Gold has four reporters that allow for an additional stringent screening.

Bait vector (pAS2-1)

There are several bait plasmids such as pAS2-1, pVA3-1 and pLAM5-1. In this study vector pAS2-1 was used to generate fusions of the target and bait, proteins with GAL DNA-BD will be expressed in yeast. The bait plasmid pAS2-1 is under the control of the full-length ADH1 promoter, has a high level of protein expression and allows selection of successfully transformed yeast cells in tryptophan-depleted media due to the presence of the*TPR1* gene (Figure 2.2).

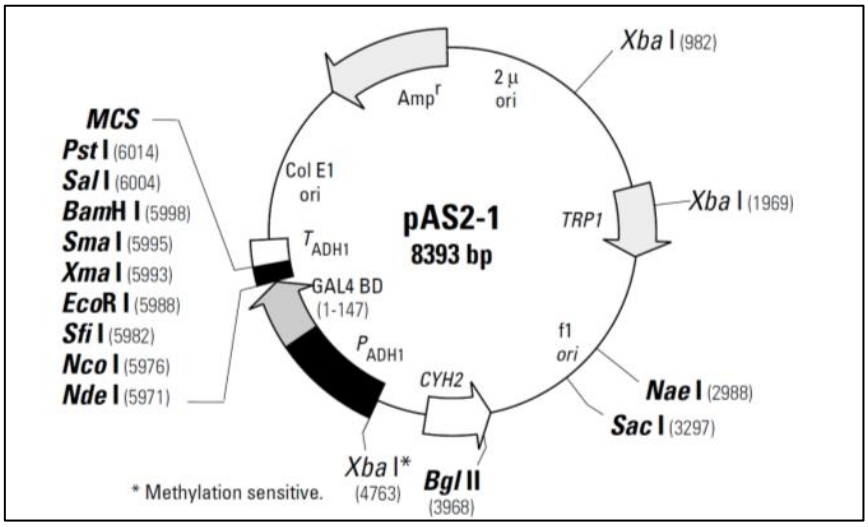


Figure 2.2: pAS2-1 map (bait plasmid; GAL4DNA binding domain) allows selection of positive cells on trp- restricted plate (Clontech Laboratories,1997).

Prey vector (pGADT7)

In this study, a cDNA library has been cloned directionally into yeast cells (strains AH109 and GOLD) using the pGADT7 prey plasmid. The target proteins are fused to the AD, which can be selected using leucine-depleted medium (due to the presence of the *LEU2* gene) when the prey plasmid is transformed into yeast cells (Figure 2.3).

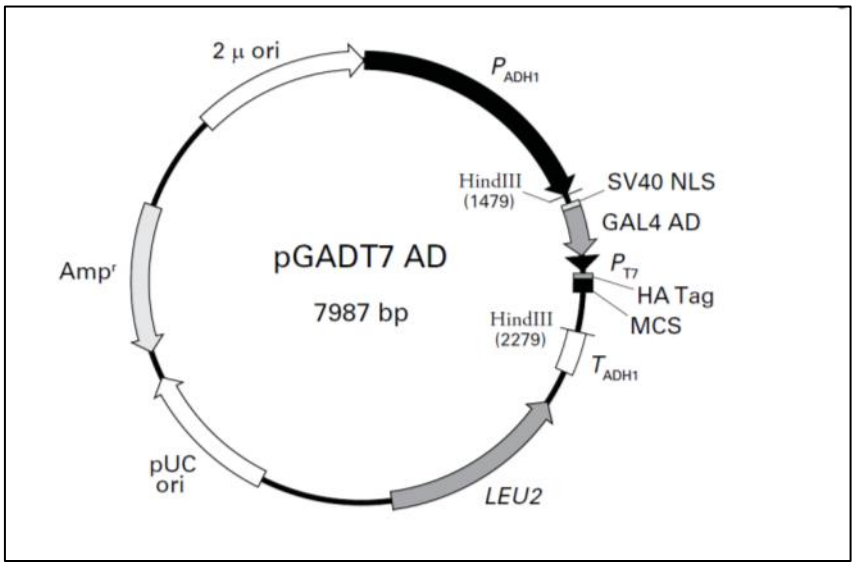


Figure 2.3: pGADT7 prey vector map (library plasmid; GAL4 activation domain), *allows nutritional selection in the leucine restricted environment* (*Clontech Laboratories, 2012).*

2.1.3. Sequence analysis of target proteins

The identification of a target protein starts with finding statistically significant matches to known protein or nucleotide sequences. In this way it is possible to infer a biological role for a gene or protein, as well as to decide whether two sequences are homologous (Nesvizhskii, 2012; Serang, Noble, 2012). The Basic Local Alignment Search Tool (BLAST) is a sequence comparison algorithm that performs local alignments of the query sequence with those in a sequence database of interest (Altschul *et al.*, 1997). It is a fast and sensitive approach that prioritizes an alignment by filtering out low complexity regions in the database (Koonin, 2003). The *E*-value illustrates the unintentional background noise in matching sequences. It is a parameter that expresses the hit numbers that are seen by chance in a known database. The *E*-value decreases as the score (*S*) between two matching sequences increases. Therefore, there is a negative exponential relationship

between *E*-value and *S* (Andrade *et al.*, 1999; Rigden *et al.*, 2011). Characterization of a novel protein, however, is limited by using BLAST (Bergman, Quesneville, 2007; Looso *et al.*, 2010).

Known protein structures in the Protein Data Base (PDB) can be searched using the query target sequence. There are a number of sequence-based comparative methods to identify proteins such as pair wise sequence alignments (Altschul *et al.*, 1994; Levitt, Gerstein, 1998; Pearson, 1996; Schuler, 1998). Sequence-based methods are improved by using Position-Specific Iterative-BLAST (PSI-BLAST) (Skolnick, Fetrow, 2000). PSI-BLAST has the potential to assign a functional role to an uncharacterized hypothetical protein and allows prediction of a structural model. Inspection of the alignment, generated by PSI-BLAST, with a protein of known structure can be used to get an indication of the structural and biochemical properties of the uncharacterized protein (Elofsson, 2002; Fischer *et al.*, 2011).

'Threading'

There are more efficient methods based on testing the compatibility of a query sequence and matching its structure in the database such as 'threading' (Lu, Skolnick, 2001; Torda, 1997). It uses sequence-structure suitability functions, such as residue-level statistical capacity functions to evaluate a sequence-structure match (Bowie *et al.*, 1991; Jones *et al.*, 1992; Sippl, Flöckner, 1996).

Protein **H**omology/analog**y** **R**ecognition **E**ngine (PHYRE) is one of the comparison modeling methods based on sequence-structure matching. Like other protein threading techniques, Phyre is capable of frequently making reliable protein models while other commonly used techniques such as PSI-BLAST cannot (Kelley, Sternberg, 2009).

The query sequence is first analyzed using PSI-BLAST against a large sequence database. The generated profile is then processed by the neural network secondary structure prediction program PsiPred and the protein disorder predictor Disopred (http://bioinf.cs.ucl.ac.uk/disoprep/). The predicted alpha helix/beta sheet locations and disordered regions are indicated graphically with color-coded confidence bars. Disordered regions (DRs) are defined as complete proteins or regions of proteins

that lack a stable tertiary structure which applies to the protein backbone instead of the residue side-chains (Figure 2.4) (Mészáros *et al.*, 2009; Schaefer *et al.*, 2010).

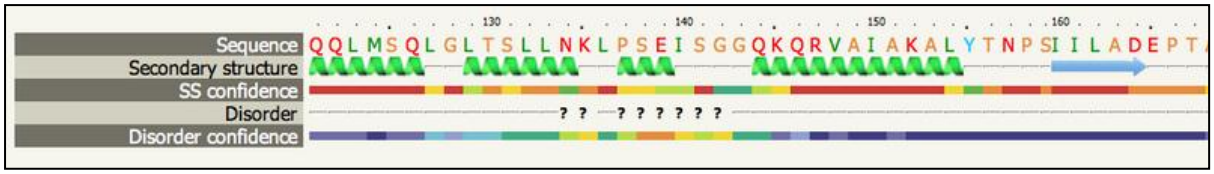


Figure 2.4: Example of a Phyre output for secondary structure and disordered prediction. *The disordered region is shown with question marks.*

Most proteins contain multiple protein domains (Björklund *et al.*, 2006). In this regard, PHYRE offers a table of matches that are color-coded according to the confidence of the prediction, which indicates the region of the query sequence matched. Therefore, it is possible to determine the domain composition of a protein.

The main results table in PHYRE offers confidence assessments, images and links to the three-dimensional predicted models and information derived from either the Structural Classification of Proteins database (SCOP) or the PDB (Figure 2.5). For each match, a link takes the user to a detailed view of the alignment between the query sequence and the template sequence of known three-dimensional structure.

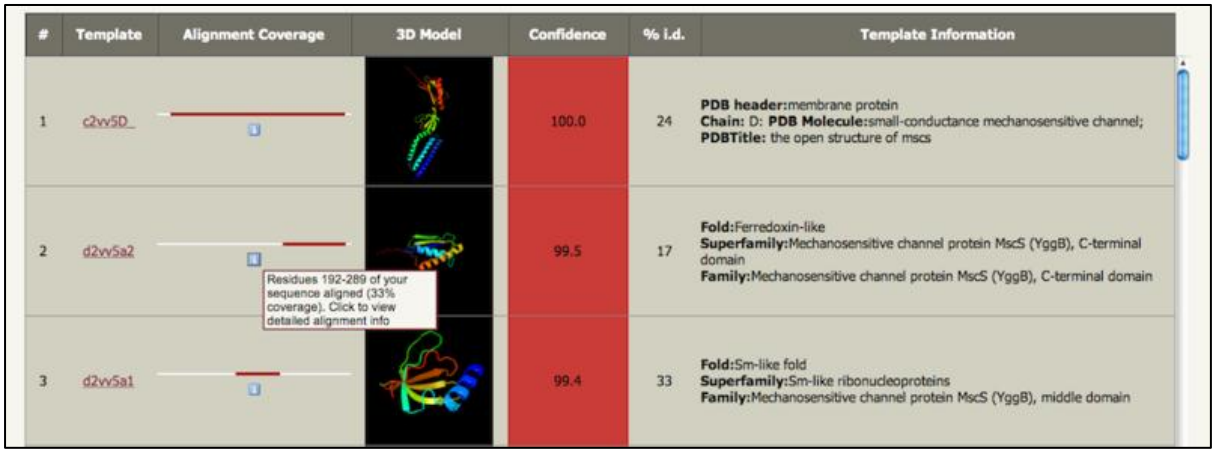


Figure 2.5: Example of a Phyre detailed template information table (the main results table). *It can provide 100 % confidence results even if the targeted sequence has only 17 % identity with the data searched (Figure taken from: www.sbg.bio.ic.ac.uk/phyre2/html [2012, September 15]).*

Every aligned residue from the main results table can be examined individually. Alignment view provides matches between predicted and known secondary structure elements and the ability to toggle information regarding patterns of sequence conservation and secondary structure confidence. In addition, Jmol or Rasmol can be used for interactive 3D viewing of the protein model (Figure 2.6).

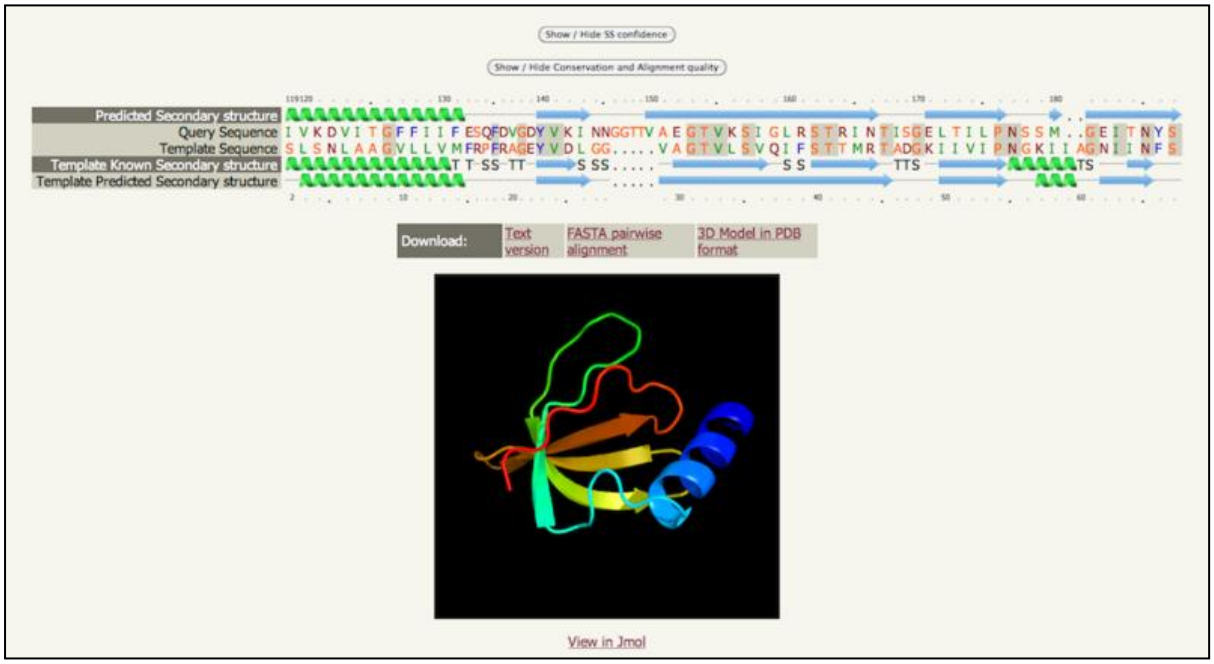


Figure 2.6: Example of a Phyre[2] detailed view of the alignment between a query sequence and known protein structure *(Figure taken from: www.sbg.bio.ic.ac.uk/phyre2/html [2012, September 15]).*

2.1.4. Protein-protein interaction confirmation

Proteins bind to other proteins in order to perform their biological functions (Chothia, Janin, 1975). Most of the important molecular processes in the cell, such as DNA replication, can only take place as a result of protein–protein interactions (De Las Rivas, Fontanillo, 2010; Deng *et al.*, 2002). Proteins bind together at their functional domains and assemble into large molecular machines that are maintained by protein-protein interactions (Hu *et al.*, 2011; Janin, Chothia, 1990; Pazos, Bang, 2006). Protein interactions have been studied in signal transduction and other metabolic or genetic/epigenetic networks (Schwikowski *et al.*, 2000). Indeed, single

protein-protein interaction can be the core of the entire interactomic system of any living cell (Kohn, 1999).

Isolation of interacting proteins using immunoprecipitation

Immunoprecipitation (IP) is a method for precipitating proteins out of solution using an antibody that specifically binds to the protein in question (Phizicky, Fields, 1995; Rigaut *et al.*, 1999). This form of affinity purification can be used to isolate and concentrate a protein of interest from a sample containing a mixture of thousands of different proteins. Immunoprecipitation requires the antibody to be coupled to a solid substrate at some point in the procedure (Harlow, Lane, 2006). This method enables the identification of expressed proteins, determination of their post-translational modifications and molecular weight (Klenova *et al.*, 2002; Xia *et al.*, 2013). More commonly, IP is used to obtain protein-binding molecules in studies investigating protein-protein and protein-nucleic acid interactions.

The precipitated protein(s) can be analysed using gel electrophoresis, mass spectrometry or Western blotting. There are several protocols which can be used to isolate proteins with immunoprecipitation such as the use of agarose-bound protein A/G (Figures 3.9 and 3.10) (Bonifacino J. S. C., 2001). The basic principle thereof is that agarose-bound protein A/G binds to the antibodies obtained from antiserum. The protein of interest (target antigen) can then be captured by the antibody on the agarose bead and can be harvested by centrifugation.

2.2. HYPOTHESES AND AIMS

Hypotheses

- Bm86 and BmATAQ interacting partners can be identified by means of the Yeast Two-Hybrid system.
- Annotation of binding partners will offer insight into the biological functions of Bm86 and BmATAQ.

Aims

- To co-transform yeast strains AH109 and GOLD with bait and prey plasmids
- To select for positive clones using nutrient-deficient media
- To sequence and identify library inserts from positive clones
- To annotate interacting proteins, improving the existing understanding of the biological roles of Bm86 and BmATAQ
- To annotate Bm86 and BmATAQ binding proteins using PSI-BLAST and threading.
- To confirm these binding partners via immunoprecipitation, SDS-PAGE and Western blot.
- To verify the amino acid sequences of the putative binding partners using LC-MS/MS.

2.3. MATERIALS AND METHODS

2.3.1. Ticks

Different life stages (nymphs and adults) of *R. microplus* ticks were obtained from ClinVet international (Pty) Ltd, Bloemfontein, South Africa, for the construction of the cDNA library. The cDNA library was constructed from whole ticks, nymphs and adults (Mariette Botha, M.Sc in Biochemistry, University of Pretoria). Construction of the bait plasmids (Bm86 and BmATAQ) were done by Rween Kalisharan at University of Pretoria. The primers used to amplify the bait and prey inserts are listed on the Table 2.2.

Table 2.2: Primers used to amplify the prey (library- first 4 primers) and bait (Bm86 and BmATAQ- last 4 primers).

Primer	Sequence	Application
5' amplimerGADT7	5'-CTATTCGATGATGAAGATACCCCACCAAACCC-3'	PCR screen of pGADT7 plasmid
3' amplimerGADT7	5'-GTGAACTTGCGGGGTTTTTCAGTATCTACGATT-3'	PCR screen of pGADT7 plasmid
Nested 5' pGADT7	5'-ATGAACATGGAGGCCAGTGAA-3'	Nested PCR screen of pGADT7
Nested 3' pGADT7	5'-CAGCTCGAGCTCGATGGATCC-3'	Nested PCR screen of pGADT7
Bm86 F	5'- GGAATTCCATATGTCATCTGTTTGCTCTGACTTC -3'	Bm86 1st PCR with restriction enzyme part (Ndel)
Bm86 R	5'- CCGGAATTCTGACTTTCCAGGATCTGGATC -3'	Bm86 1st PCR with restriction enzyme part (EcoRI)
BmATAQF	5'- CGGGATCCCCGAAGAGATTCCCGATATGGAC -3'	1st PCR forward primer with restriction enzyme part (BamH1)
BmATAQR	5'- TTCCGCGGCCGCTATGGCCGACGTCGACCA CGACTGGTTTGTCTTTGGGG -3'	1st PCR reverse primer with restriction enzyme part (Sal1)

2.3.2. Small scale yeast transformation

Yeast cells were transformed with pAS2_1_Bm86 (bait) using the Frozen-EZ Yeast Transformation II™ kit as per manufacturers' guidelines (Zymo Research Corporation, USA). Native yeast cells were allowed to grow to mid-log phase (OD_{600} ~0.5, 5 x 10^7 cells/ml), to ensure maximal transformation efficiency with shaking at 30°C in YPDA media. Transformants (50-100 µl) were plated on tryptophan depleted SD minimal media and incubated at 30°C for 3-6 days upside down.

2.3.3. Test for autonomous reporter gene activation

Yeast cells transformed with bait plasmid were tested for autonomous transcriptional activation by plating the cells on SD/-Trp (synthetic drop-out), SD/-His/-Trp (double drop-out, DDO) and SD/-His/-Trp/-Leu (triple drop-out, TDO) media and incubation at 30 ºC for 3-6 days.

2.3.4. Sequential library-scale transformation of yeast cells

Gold yeast cells containing the BmATAQ DNA binding (bait) construct were inoculated in 50 ml SD/-Trp and grown overnight with shaking at 30°C. After cell titer determination (7.2 x 10^8 cells), yeast cells were collected by centrifugation (3000 x g, 5-10 min) and diluted in 150 ml YPDA media. Once the titer had reached 2 x 10^7 cells per ml, cells were collected by centrifugation (3000 x g, 5-10 min) following incubation at 30°C with shaking for 4-5 hours. Sequential library-scale and co-transformation of the yeast cells with the DNA-AD/Prey plasmid were performed using the Frozen-EZ Yeast Transformation II™ kit. Cells were transformed using 1.5 µg library plasmid. Both Gold and AH109 yeast cells were plated on DDO (-Leu/-Trp) plates (15 cm in diameter) and incubated at 30°C for 4-6 days.

2.3.5. Two-hybrid screening of reporter genes

Co-transformed cells were scraped from the DDO plates into 2 ml SD/-Leu/-Trp and vortexed briefly, after which 1 ml of resuspended yeast cells were plated onto SD/-Leu/-Trp/-His (TDO, triple drop-out) plates and incubated at 30°C for 5-10 days until colonies started to appear. The remaining 1 ml co-transformed cells were stored at -70°C in DDO medium containing 25% glycerol. Colonies grown on TDO were transferred to two TDO master plates (duplicates) with sterile wooden sticks and incubated overnight at 30°C. Positive colonies from one master plate were transferred onto a SD/-Leu/-Trp/-His/-ade (QDO, quadruple drop-out) plate and incubated overnight at 30°C while the other master plate was stored at 4°C. Positive clones from QDO plates were subjected to nested colony PCR and colony-lift β-galactosidase filter assays.

2.3.6. Colony-lift β-galactosidase filter assay

A sterile dry filter paper was placed over the QDO plate surface and gently rubbed to help colonies cling to the filter paper. When the filter paper got evenly wet, it was removed from the plate and incubated in liquid nitrogen, colonies facing up, for 20 seconds and thawed at room temperature. This procedure was repeated 4 times to permeabilize the cells. Subsequently, the filter with positive colonies was incubated between 2 filter papers pre-soaked in 100 ml Z buffer (16,1 g/L $Na_2HPO_4.7H_2O$; 5,50 g/L $NaH_2PO4.H_2O$; 0,75 g/L KCl; 0,46 g/L $MgSO_47H_2O$; pH 7.0) containing 167 µl of X-gal solution (20 mg/ml) and incubated at room temperature until a blue color appeared (3-5 hours) indicating promoter activity. The assay containing the positive control (yeast cells Gold/AH109 with the control plasmid pCL1) turned blue in about 1.5 hours.

2.3.7. Nested Colony-PCR screening of positive colonies

Positive colonies from the QDO plates were analyzed using PCR. They were briefly touched with a tip and the cells at the edge of the tip transferred into a PCR reaction containing 12,5 µl of (2X) KAPA ready mix (250 rxn, KAPA Biosystems; USA),10 mM

primers and made up to a final reaction volume of 25 µl with PCR water. Amplification was performed in a Applied Biosystems GeneAmp ® PCR System 2700, USA, with an incubation at 94 ºC for 7 min to disrupt the cell membrane and 30 cycles of denaturation at 94°C for 30 seconds, annealing at 58°C for 30 seconds, and extension at 72°C for 2 min. All the conditions were kept the same for the nested PCR using the nested pGADT7 primers (Table2.2).

2.3.8. Plasmid isolation

Plasmid isolation from yeast cells was done in accordance with the method of Dr. Moolman-Smook (Department of Medical Biochemistry, University of Stellenbosch, South Africa). Co-transformed yeast cells were grown overnight in 1 ml minimal media (SD/-Trp/-Leu, DDO) at 30ºC with shaking. The culture was then incubated in extra 4 ml YPDA media for an extra 4-5 hours at 30ºC with shaking. The cells were collected by centrifugation (3000 x *g* for 6-10 minutes) and resuspended in 200 µl Smash & Grab buffer (1% sodium dodecyl sulfate-SDS, 2% Triton X-100, 100 mM NaCl. 10 mM Tris-HCl and 1 mM ethylenediaminetetraacetic acid at pH 8. Resuspended cells were lyzed via the addition of 100 µl glass beads (425-600 micron), 200 µl phenol: chloroform: isoamylalcohol (25:24:1) and vortexing for 5 min. The sample was then centrifuged at 16045,79 x g for 5 min and the aqueous layer was transferred to a new eppendorf tube. The plasmid was precipitated by adding 0.5 volumes of ammonium acetate (7.5 M, pH 5) and 2 volumes of 100% ethanol and centrifugation (16045,79 x g for 25 min). The resultant pellet was washed with 70 % ethanol, centrifuged (16045,79 x g for 25 minutes) and dried in vacuum. The dried pellet was finally resuspended in 20 µl sterile water.

2.3.9. Library plasmid isolation via transforming into KC8 *E. coli* strain

After yeast two-hybrid screening, positive prey plasmids need to be rescued for further analysis. The *E. coli* strain KC8 has mutations in its *trpC*, *leuB*, and *hisB* genes that complement the yeast TRP1, LEU2 and HIS3 proteins. KC8 cells were electroporated with the plasmids isolated from yeast cells. The positive clones were selected on leucine-deficient M9 minimal media (1.28% Na_2HPO_4-$7H_2O$, 0.3%

KH_2PO_4, 0.05% NaCl, 0.1% NH_4Cl, 0.2% of 1 M $MgSO_4$-$7H_2O$, 0.01% ml of 1 M $CaCl_2$ and 0.2% glucose) containing 0.1% of 50 mg/ml ampicillin. Library plasmids were isolated from KC8 cells using the Zyppy™ Plasmid Miniprep kit (Zymo Research Corporation, USA). Plasmids were subsequently sequenced by Inqaba Biotechnical Industries (Pty) Ltd, South Africa using plasmid-specific (pGADT7) primers (Table 2.2).

2.3.10. Data analysis

Nucleic acid and protein sequences obtained were manually inspected using BioEdit (www.mbio.ncsu.edu/BioEdit/bioedit.html) and GeneDoc (www.nrbsc.org/gfx/genedoc/) sequence alignment editor software. Sequence analysis was performed using Position Specific Iterated Basic Local Alignment Search Tool (PSI-BLAST) (http://www.ncbi.nlm.nih.gov/BLAST).Sequences from PSI-BLAST analysis were subjected to threading by using the Protein Homology/analogy Recognition Engine V 2.0 (Phyre2), a three dimensional protein structure prediction program (http://www.sbg.bio.ic.ac.uk /phyre2/html/). Sequence alignments were done using Clustal-W (www.ebi.ac.uk/clustalw/) and Mafft (www.ebi.ac.uk/mafft/).

2.3.11. Tick gut homogenate preparation

Guts of five 20-day fed female *R. microplus* ticks were suspended in homogenization buffer (20 mM Tris, 150 mM NaCl, 0.05% Tween 20 and 0.37 mg/ml complete proteinase inhibitor). The mixture was physical shearing using needles (20G) and homogenized at maximum speed using a Heidolph RZR2050 tissue homogenizer (Wotol, Europe), followed by sonification (2 s pulse, duty cycle 30/40) for 2 seconds at 3W output using a Branson Sonifier 450 (Instrulab, RSA). The homogenate was centrifuged at 4°C, 24 800 x g for 30 minutes to separate the soluble proteins (supernatant) from membrane-bound proteins. After removing the supernatant, the pellet was resuspended in Tris-buffered saline (TBS) (20 mM Tris, 150 mM NaCl, 0.05% Tween 20; pH 7.4) containing 0.5% CHAPS, for one hour at 4°C. The

homogenates were either immediately subjected to immunoprecipitation or stored at -20°C.

2.3.12. Immunoprecipitation (pull-down assay) of the binding partners

Agarose-bound protein A (100 mg) was allowed to swell in 1 ml TBS wash buffer (20 mM Tris, 150 mM NaCl, 0.05% Tween 20 at pH 7.4) by incubation at room temperature for 15 minutes on a shaking platform. This procedure was repeated three times. The wash buffer was removed and 800 µl fresh wash buffer were added containing 250 µl α-Bm86 or α-BmATAQ antisera. The mixture was placed on a rotating mixer for 3 hours at 4°C and incubated for 5 minutes at room temperature to allow the agarose-protein-A-antibody complex to settle at the bottom of the tube, after which the supernatant was carefully removed. The pellet was washed three times by adding 1 ml wash buffer and rotating the suspension for a further 10 minutes at 4°C at a slow speed.

In order to isolate the interacting partners, 500 µl wash buffer, 500 µl gut proteins (7 mg/ml) and 4 mg/ml recombinant Bm86/BmATAQ were added to the agarose-protein-A-antibody complex. The reaction was incubated overnight at 4°C in a rotating mixer to allow the interacting partners to bind to the target proteins. The mixture was then washed 3 times with100 µl glycine/elution buffer (0.1 mM glycine; 0.5 NaCl; pH 2.4) and the samples were put in a rotating mixer at 4°C for 5 minutes. The supernatant was removed and put into a 0.6 ml tube containing 20 µl Tris (2 M Tris-HCl, pH 9). The proteins were either directly visualized using SDS-PAGE or stored at -20 °C.

2.3.13. SDS-PAGE

Proteins obtained from the immune-precipitation were mixed (1:1 ratio) with sample buffer (2% SDS, 0.1% glycerol, 0.05 M Tris-HCl, 0.025 % bromophenol blue and 0.05% β-mercaptoethanol pH 6.8) and boiled for 5 minutes prior to loading a 4% stacking gel and a 12% separating polyacrylamide gel prepared according to the method from Laemmli (1970). Gels were polymerized by the addition of 0.05%

ammonium persulfate and 0.001% N, N, N′,N′-tetramethylethylenediamine (TEMED). Samples were run on the BAYGENE electrophoresis system (B6-Power 300, China) at 6 V/cm until the separating gel was reached and then continued at 10 V/cm for 5 hours. PageRuler™ Unstained Protein Ladder or PageRuler™ prestained protein ladder was used to determine the molecular weight of the proteins.

Tricine SDS-PAGE

Tricine SDS-PAGE is able to separate proteins with molecular weight ranging from 1 to 100 kDa (Schagger, Von Jagow, 1987). The 4% stacking gel consists of 8% acrylamide stock solution (49.5% acrylamide and 3% bisacrylamide) and 24.8% gel buffer (3M Tris-HCl; 0.3% SDS; pH 8.4) The 16% separating gel consists of 33.4% acrylamide stock solution and 33.4% gel buffer. Both the stacking and separating gels were polymerized by the addition of 0.05 ammonium persulfate and 0.001% TEMED. Protein samples were prepared as described previously and ran on the prepared tricine gel in cathode buffer (0.1 M Tris; 0.1 M Tricine; 0.1% SDS; pH ~8.25) and an anode buffer (0.2 M Tris-HCl; pH ~8.9) initially at 3 V/cm for 1 hour to stack the proteins then at 10 V/cm for 7 hours.

2.3.14. SDS-PAGE staining

SDS-PAGE gels were stained using silver or colloidal Coomassie Blue (CCB) (Merril *et al*. 1981, Nirmalan, 2004). For silver staining, proteins were fixed overnight by incubation in 45% (v/v) methanol and 5% (v/v) acetic acid, followed by sensitizing for 2 min in 0.02% (w/v) sodium thiosulfate, and two rinsing steps with $ddd H_2O$. Subsequently, 200 ml ice cold 0.1% (w/v) silver nitrate was added and gels were incubated at 4°C for 30 minutes and rinsed twice with $ddd H_2O$. Development was initiated by adding a freshly prepared solution containing 2% (w/v) sodium carbonate and 0.004% (v/v) formaldehyde. The reaction was terminated with the addition of 1% (v/v) acetic acid. Silver stained polyacrylamide gels were subsequently stored at 4°C.

Staining with CCB was performed using 2% colloidal Coomassie brilliant blue G250 stock solution (v/v) phosphoric acid, 10 % (w/v) ammonium sulfate and 0.1% (v/v) Coomassie brilliant blue G250 mixed 4:1 with methanol. Polyacrylamide gels were immersed in the colloidal Coomassie solution and put on a shaking platform overnight, and then rinsed with a solution containing 25% (v/v) methanol and 10% (v/v) acetic acid. The gels destained via incubation in 25% (v/v) methanol until the background became clear. Gels were visualized using the Bio-RAD molecular Imager® Gel Doc™ XR+ imaging system (Bio-RAD, USA) and stored at 4°C in 1% (v/v) acetic acid until use for mass spectrometry.

2.3.15. Western blot analysis

Following SDS-PAGE, proteins were immediately transferred onto a polyvinylidene fluoride (PVDF) membrane at 10 V for 30 minutes using the semi-dry Trans-Blot electrophoretic transfer cell system (Bio-Rad, USA). The PVDF membrane was first activated by soaking in 100% methanol for 5 seconds and washed in 10 mM CAPS (3-(cyclohexylamino)-1-propanesulfonic acid, pH 9) for 5 seconds. The blotted membranes were placed in sealed plastic bags and incubated overnight in TBS blocking buffer (20 mM Tris-HCl, 150 mM NaCl, 5% skim milk, pH 7.4) at 4°C. The membranes were washed three times with wash buffer (20 mM Tris, 150 mM NaCl, 0.05 Tween 20; pH 7.4) for 10-15 minutes and incubated with 1:5000 diluted primary antibody (anti-RDH14 antibody produced in rabbit) for 3 hours with agitation. This was followed by three washing steps and incubation with 1:5000 diluted secondary antibody (anti-rabbit H+L chain) for 1 hour.

The membranes were finally washed three more times with wash buffer and the protein-antibody complexes visualized by using SuperSignal® chemiluminescent substrate (Pierce-Thermo Scientific, USA). Shortly, protein-antibody complexes on the membranes were incubated with equal volumes (4 ml) of SuperSignal West Pico Luminol/Enhancer and Stable Peroxidase solutions for 5 min. The membranes were then exposed to Hyperfilm High Performance chemiluminescent film (Amersham Biosciences, England) for 3-30 minutes. The films were developed using the HiPam universal developer (IIIford, USA) and fixed in HiPam universal fixer (IIIford, USA). Membranes were briefly washed in dddH_2O and left to dry.

2.3.16. LC-MS/MS

Mass spectrometry analysis of proteins was done by Dr. Salome Smit (Central Analytical Facilities, Proteomics Laboratory, Stellenbosch University). Briefly, gel pieces were washed with water, 50% acetonitrile and 50 mM ammonium bicarbonate. The proteins were then trypsinized overnight at 37°C. The resultant peptides were extracted via incubation with 70% acetonitrile in 0.1% formic acid for 30 minutes. Peptides were dried and stored at -20°C and dissolved in 5% acetonitrile and 0.1% formic acid before performing nano-LC chromatography.

For liquid chromatography, separation of peptides was performed using an EASY-Column with a flow rate 300 nl / min. The gradient used was from 5-40% solvent B (100 % acetonitrile in 0.1% formic acid) in 20 min, 40-80% solvent B in 5 min and kept at 80% solvent B for 10 min. Mass spectrometry conditions were 1.5 kV, capillary temperature of 200 °C, with no sheath and auxiliary gas flow. The ion selection threshold was 500 counts, with activation Q-value of 0.25 and activation time of 10 ms for MS/MS.

Proteins were identified using automated database searching (Mascot, Matrix Science, London, UK) of all tandem mass spectra against the tick database (BmGI) as well as non-redundant database and finally against Swiss-Prot (web.expasy.org/docs/swiss-prot_guideline.html). A Mascot score with a *P* value smaller than 0.05 as determined by Proteome Discoverer 1.3 was used. Percolator was also used for validation of search results. In Percolator a bait database was searched with a strict false discovery rate (FDR) of 0.02 and a relaxed FDR of 0.05 with validation based on the q-value (Central Analytic Facilities, MS unit, Stellenbosch University).

2.4 RESULTS AND DISCUSSION

2.4.1. Transformation of yeast cells with bait/GAL4 BD

As the bait plasmid pAS2-1 contains the *TRP1* reporter gene, yeast cells transformed with bait plasmid were plated on SD/-Trp media in a small-scale transformation of yeast and grew successfully. Analysis of autonomous reporter gene activation showed that the yeast cells were suitable to use for the bait plasmid screening on the appropriate minimal media as no growth was seen on the TDO and QDO plates after incubation at 30°C for 4-6 days.

2.4.2. Transformation of yeast cells with the prey/library plasmid construct

As the *LEU2* gene is found in the library plasmid, the positive yeast clones co-transformation with both bait (pAS2-1) and the library (pGADT7) plasmids were selected for on SD/-Trp/-Leu (DDO) plates. After 4 days of incubation at 30°C, clones were visible. Incubation was prolonged for sufficient yeast growth, in order to continue with yeast two-hybrid screening. Positive clones from the DDO plates were plated onto TDO and QDO plates to identify interacting partners (Figure 2.7) Interactions between the product of bait and prey plasmid inserts were detected as cell growth was obtained on both TDO and QDO media. From 96 putative BmATAQ interacting clones on the TDO master plate, only 68 colonies grew on a QDO media (Figure 2.7 A and B). Initial two-hybrid screening for potential Bm86 interacting partners on TDO resulted in the growth of 100 colonies, however, only 14 of these were able to grow on a QDO (Figure 2.7 C and D). The remaining colonies, not able to grow on QDO, developed a reddish-pink color which might be an indication of adenine depletion (Matecic *et al.*, 2010).

The *ADE2* gene could therefore not be expressed at high enough levels as a result of a weak interaction between bait and prey plasmid insert products in these colonies. Consequently, colonies that did grow without color change on QDO plates were used for further analysis (Figure 2.7).

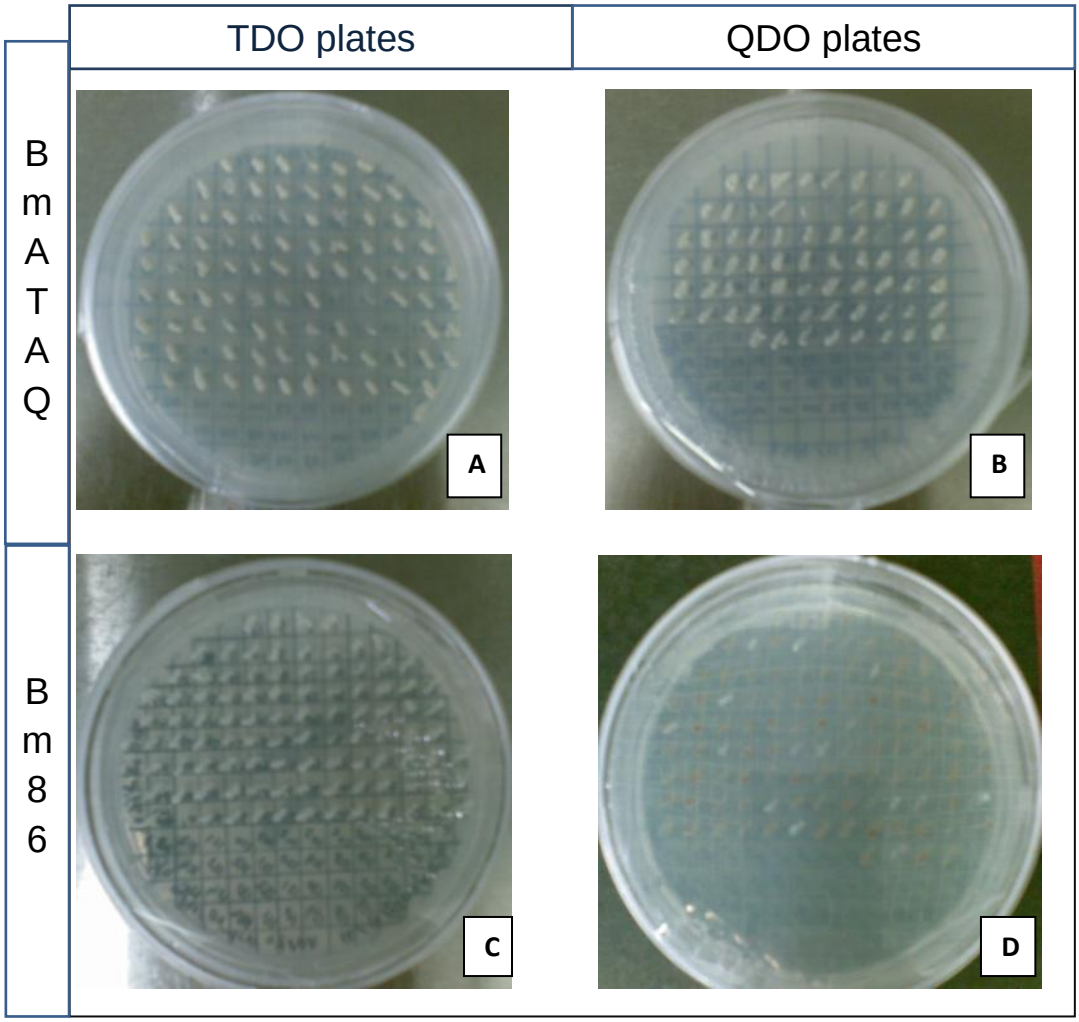


Figure 2.7: Growth of yeast cells on TDO and QDO media which are co-transformed with BmATAQ (A and B) and Bm86 (C and D). *From BmATAQ TDO plate only 68 of 96 colonies showed growth on QDO which were used for further analysis. From Bm86 TDO plate only 14 of 100 colonies appeared to survive on QDO plate, showing putative interactions between bait and prey. Most of the cells turned into a pink colour on QDO (B and D) media after they were transferred from TDO (A and C) plate as an indication of weak interaction or heavy stress on cells such as depletion of adenine.*

2.4.3. Nested colony PCR screening of positive clones

The presence and size of the library plasmid inserts were determined using nested colony PCR on QDO positive clones. Unique clones were identified by restriction enzyme digestion of BmATAQ library inserts with *Eco*RI and *Bam*HI. Eight distinct groups of clones were identified according to their size different fragment sizes obtained after digestion (Figure 2.8). In total 64 clones were screened with nested colony PCR and 36 of these were classified into 8 unique groups (Figure 2.8, Table 2.3).

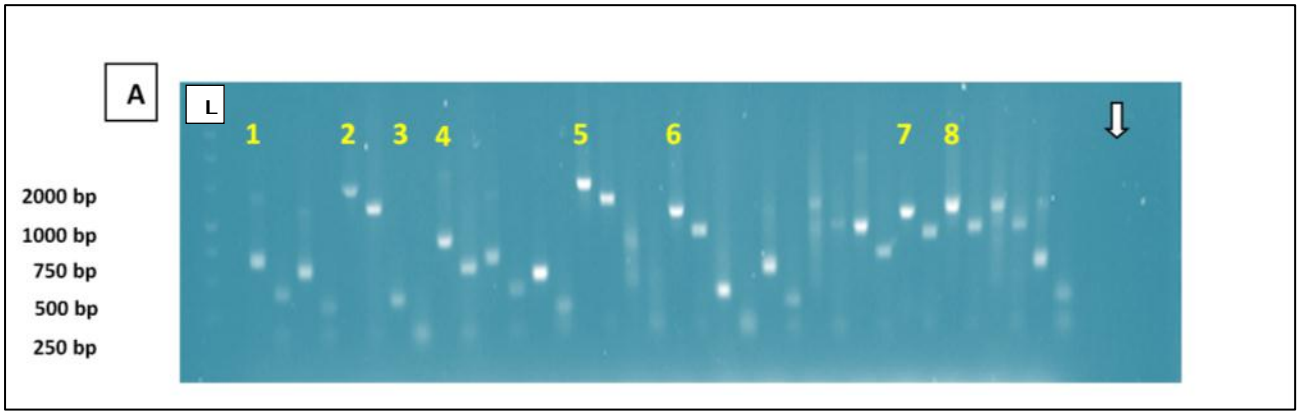


Figure 2.8: (A) Agarose gel electrophoresis after colony PCR and *Eco*RI and *Bam*HI digestion of clones containing putative ATAQ binding partners. *Every numbered lane represents a unique clone and every lane after the numbered ones represent digested clone. Arrow indicates the negative control lane. L represents the pre-stained DNA Ladder.*

Table 2.3: Estimated number of unique predicted ATAQ groups, clones in each group and sizes of each group according to nested colony PCR. *In total 8 groups selected as unique according to estimated sizes, six groups had five and two groups had three clones in each with similar sizes. Overall 36 clones were analyzed with nested PCR, size of the clones were ranged between >750 bp and >2000 bp.*

	Number of clones in the group	Estimated size of the inserts
Group 1	5	>750 bp
Group 2	5	~ 2000 bp
Group 3	5	<750 bp
Group 4	5	<1000 bp
Group 5	3	>2000 bp
Group 6	5	>1000 bp
Group 7	3	~1000 bp
Group 8	5	<2000 bp

Due to the fact that only a few unique clones could be obtained for Bm86, restriction enzyme mapping was not performed. The colony PCR products were visualized by agarose gel (1%) electrophoresis and 14 unique clones were identified (Figure 2.9).

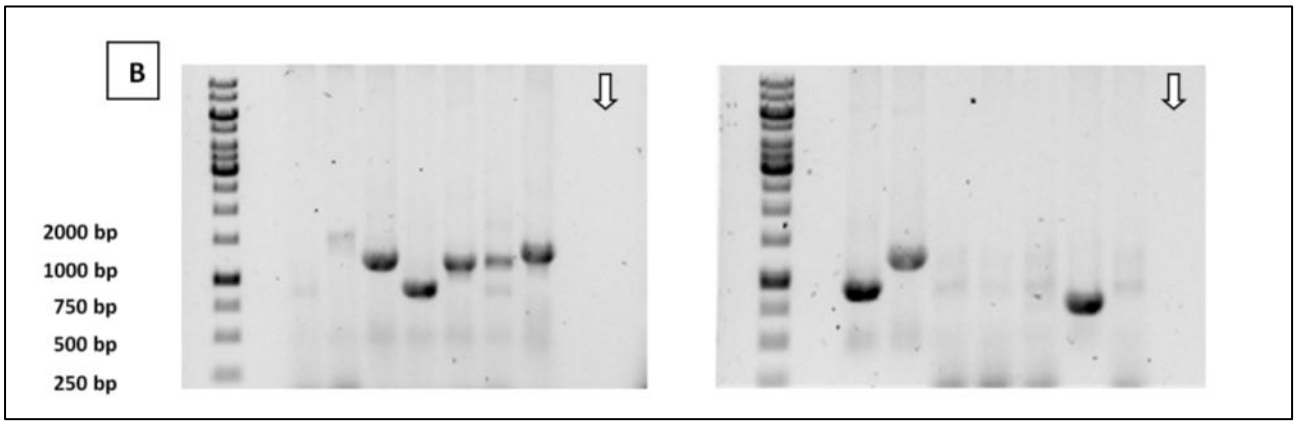


Figure 2.9: Agarose gel electrophoresis after colony PCR of putative Bm86 binding partners. *Every lane represents a different QDO positive clone. Arrows indicate the lanes with the negative control.*

Two clones from each of the eight groups of the positive library plasmids for ATAQ and all 14 positive clones for Bm86 were isolated following transformation into *E. coli* KC8 cells and plasmid sequencing using pGADT7 specific primers (Table 2.2). After sequence analysis (discussed later), selected library plasmids were used to retransform yeast cells containing the appropriate bait plasmid of interest. Co-transformed yeast cells were then grown on DDO, TDO and QDO plates to reconfirm protein interactions. Protein interactions were also further confirmed by β-galactosidase assays for both BmATAQ and Bm86.

After nested PCR analyses, 36 positive clones were identified as putative BmATAQ interacting candidates. Sequencing was done for 16 selected library inserts, two from each of the eight groups, however; only eight of the 16 clones could be sequenced. Clones 12 and 13 were chosen as they belonged to same group, with similar sizes after nested colony PCR and restriction enzyme mapping. This was supported using sequencing results, which indicated that the inserts of clone 12 and 13 were 98 % identical (Appendix 1). Hence, the identification of unique clones via restriction

enzyme mapping was a reliable method to reduce the numerous positive clones to a reasonable number for further analysis.

Due to the limited number of positive clones for putative Bm86 interacting candidates, it was not necessary to perform restriction enzyme mapping to select unique inserts. All 14 inserts were therefore subjected to sequencing, but only 4 out of 14 could be sequenced successfully. Clone 93 of the Bm86 interacting candidates has the same nucleic acid sequence as clone number 12 and 13, which could indicate that Bm86 and BmATAQ are involved in the same metabolic pathway/s. The sequence identity between clone 93 and clone 12 was 96 % (Figure 2.10).

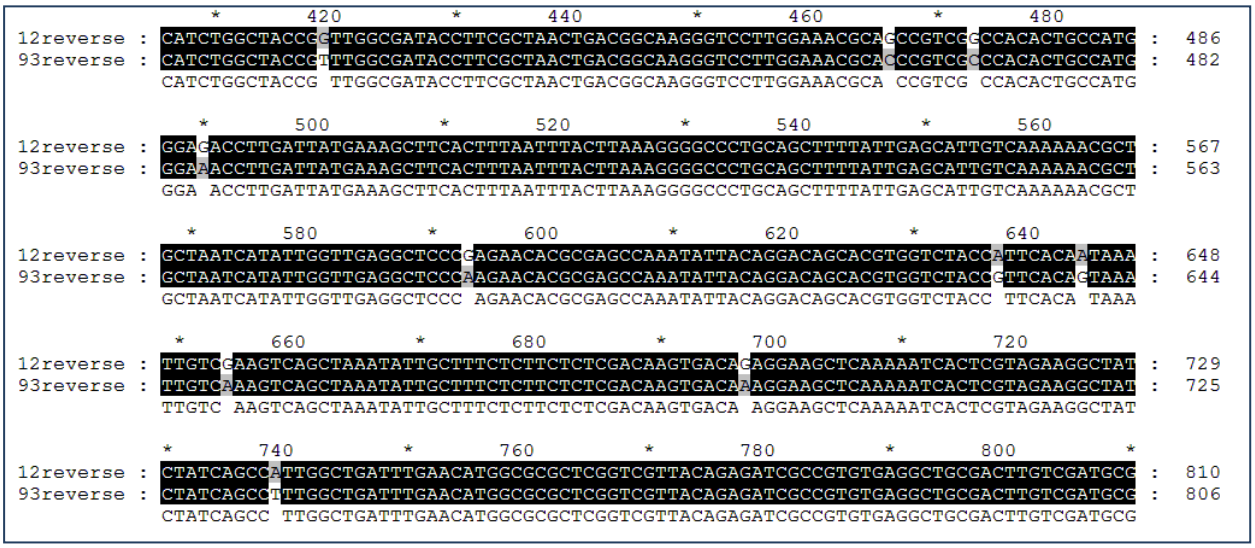


Figure 2.10: Section of clone 12 (BmATAQ) and clone 93 (Bm86) nucleotide alignment. *There are 10 nucleotide differences between clone 12 and clone 93, the identity between the two sequences is 96%.*

2.4.4. Examination of the sequence data

Prior to further analysis, vector sequences were removed from the nucleotide sequences of binding proteins before translation. The deduced amino acid sequences were then analyzed using PSI-BLAST where only two of both the Bm86 and BmATAQ interacting candidates were identified with significant results. These sequences corresponded to clones 50 and 12 and clones 53 and 93 as putative BmATAQ and Bm86 interacting partners, respectively.

PSI-BLAST analysis showed that the putative Bm86 interacting partners clone 53 and clone 93 were predicted to be a multi domain Kunitz-like protease inhibitor (E value of 5E-21), and an aldehyde dehydrogenase (ALDH) (E value of 5*E*-13), respectively. The outcomes of the PSI-BLAST were confirmed with BLAST analyses performed against the tick database (Maritz-Olivier *et al.*, 2012) and clone 53 was found to match with CK176583 (3*E*-69) that was identified to be a multi-domain Kunitz-like protease inhibitor. Clone 93 matched with TC16791 (with 60 % identity) that was predicted to be an aldehyde dehydrogenase of *Rhipicephalus pulchellus* (the zebra tick).

PSI-BLAST searches of the identified putative BmATAQ interacting partners clone 50 and clone 12 revealed that these proteins are predicted to be a retinol dehydrogenase (E value of 8*E*-50) and an aldehyde dehydrogenase (E value of 5*E*-13), respectively. BLAST analysis against the tick database confirmed these results and clone 50 was found to be identical to TC24065 (E value of 5*E*-142) predicted to be a retinol dehydrogenase of *Xenopus tropicalis* (western clawed frog). Clone 12 was detected as TC16791 (with 60 % identity) that is a predicted aldehyde dehydrogenase of *R. pulchellus.* Both of clone 12 and clone 93 corresponds to the same putative aldehyde dehydrogenase predicted for *R. pulchellus.*

BmATAQ Clone 50, putative Retinol dehydrogenase (RDH)

The amino acid alignment of the longest predicted open reading frame of clone 50 with the first best hits from the PSI-BLAST analysis showed that this clone was closely related to retinol dehydrogenases (Figure 2.11). Retinol dehydrogenase (RDH) and related proteins belong to the classical short-chain dehydrogenases/ reductases (SDR) family (Liden, Eriksson, 2006). In this protein family, RDHs are found in a SDR subgroup. Proteins in this subfamily share the glycine-rich NAD-binding motif of the classical SDRs, have a partial match to the canonical active site tetrad, but lack a typical active site serine (Kavanagh *et al.*, 2008). SDRs are a functionally diverse family of oxidoreductases that have a single domain with a structurally conserved Rossmann fold (alpha/beta folding pattern with a central β-sheet), an NADPH-binding region, and a structurally diverse C-terminal region. Characteristically, this turn exhibits a consensus binding pattern similar to GlyXGlyXXGly (X represents any amino acids), in which the first 2 glycine residues participate in NAD(P)-binding, and the third facilitates close packing of the helix to the beta-strand (Wierenga *et al.*, 1986). Typically, proteins in this family contain a second domain in addition to the NADB domain, which is responsible for specifically binding a substrate and catalyzing a particular enzymatic reaction (Liu *et al.*, 2013).

To ensure binding with the appropriate coenzyme, these oxidoreductases contain the following motif: (Gly or Ala)-(Xaa)3-(Gly, Ala or Thr)-(Ile or Leu)-(Gly, Ala or Ser)-(Xaa)10-(Gly or Asn) (Kalanon, McFadden, 2008). This motif, however, could not be identified in clone 50.

Sequence identity between different SDR enzymes typically varies from 15-30%, but they share the Rossmann fold and characteristic NAD-binding and catalytic motifs (Gatzeva-Topalova *et al.*, 2004). These enzymes catalyze a wide range of reactions including the metabolism of steroids, cofactors, carbohydrates, lipids, aromatic compounds and amino acids, as well as acting in redox sensing (Kavanagh *et al.*, 2008).

Classical SDRs have a ThrGlyXXX[AlaGly]XGly motif or an Asparagine (Asp/N) residue serves as a cofactor-binding site, and a TyrXXXLys active site motif and upstream Asp, with the Tyrosine (Tyr/Y) residue of the active site motif serving as a critical catalytic residue (Figure 2.11). The active catalytic residue Lysine (Lys/K) is present in all five organisms' RDH sequences. However, the upstream Asp residue of the active site is not present in all organisms but tick including human RDH (Liden, Eriksson, 2006).

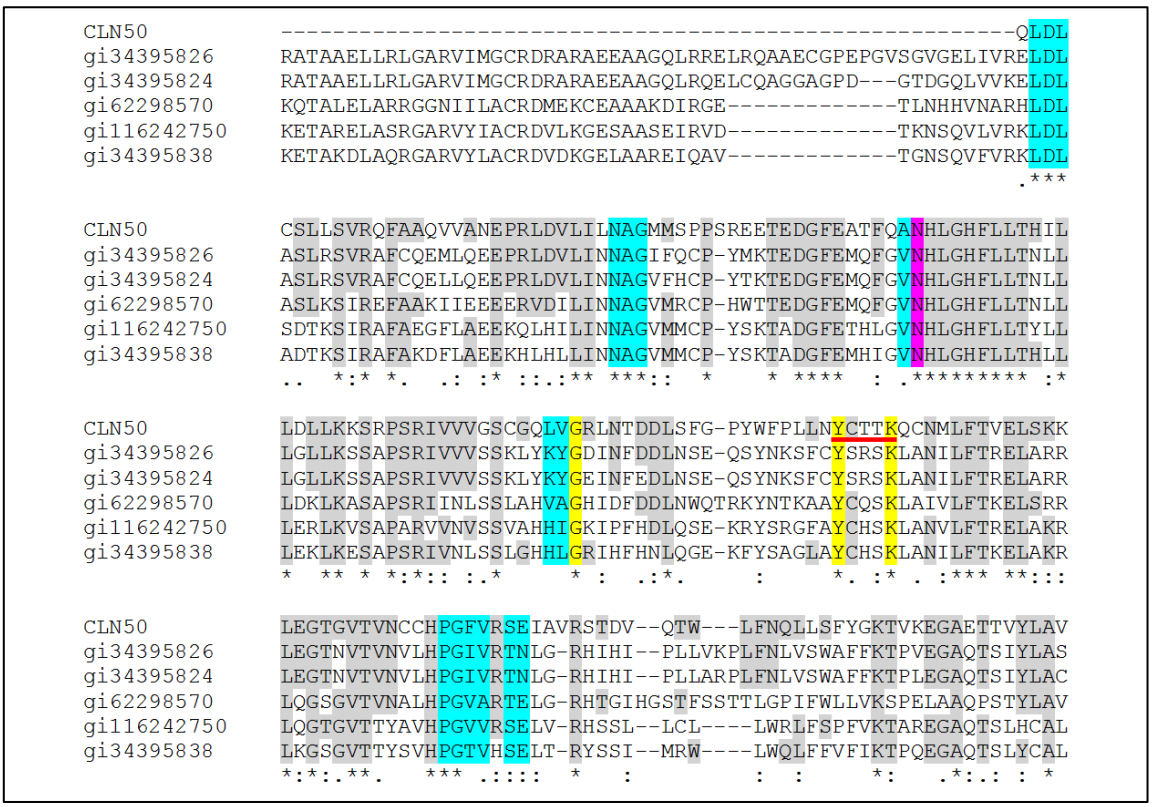


Figure 2.11: Section of amino acid alignment of putative tick RDH (clone 50) with PSI-BLAST hits. *Blue represents NAD binding sites and yellow represent active residues. The YXXXK active site motif is underlined with red. Purple represents a possible cofactor binding or a catalytic residue. Gi:34395826 : Retinol dehydrogenase 14, Homo sapiens (human), E value of 8E-50; Gi:34395824: Retinol dehydrogenase 14, Mus musculus (house mouse), E value of 1E-48; Gi:62298570: Retinol dehydrogenase 13, Homo sapiens (human), E value of 4E-46; Gi:1162422750: Retinol dehydrogenase 12, Homo sapiens (human), E value of 2E-44; Gi:34395838: Retinol dehydrogenase 11, Mus musculus (house mouse), E value of 3E-44.*

Most of the typical characteristic residues of retinol dehydrogenase were also found on the best hits from the tick database (BmGI). Possible functional residues are highlighted on the amino acid alignments of clone 50 with BmGI hits (Figure 2.12).

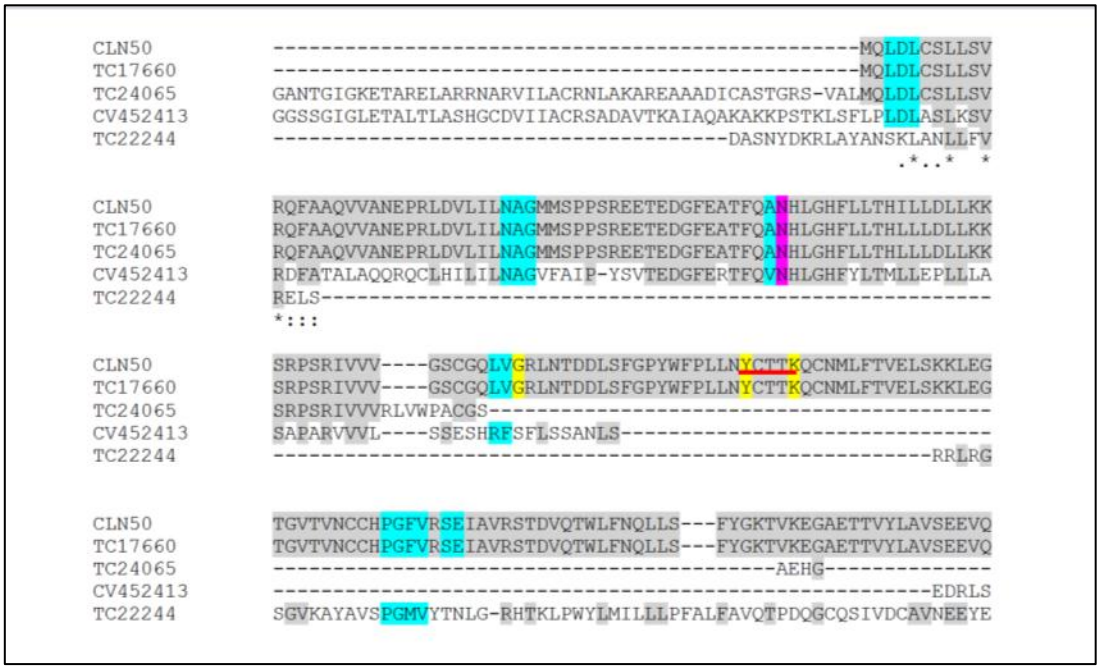


Figure 2.12: Section of amino acid alignment of predicted tick RDH (clone 50) with BMGI BLAST hits. *Blue represents NAD binding sites and yellow represent active residues. The YXXXK active site motif is underlined with red. Purple represents a possible cofactor binding or a catalytic residue.TC17660: putative retinol dehydrogenase, R. micproplus (cattle tick), E value of 5E-14.; TC24065: Retinol dehydrogenase 14 Xenopustropicalis (western clawed frog), E value of 5E-142; CV452413: Short-chain dehydrogenase Aedesaegypt (yellow fever mosquito), E value of 2E-06; TC22244: Retinol dehydrogenase 14 Xenopustropicalis (western clawed frog), E value of 5E-5.*

Structural analysis via threading using the Phyre website confirmed that the insert of clone 50 was a RDH-like protein, as it contained the NAD(P)-binding Rossmann-fold domain with 99.9% confidence. However, 25% sequence identity was obtained to the crystal structure, c3o26A, from the opium poppy, *Papaver somniferum* (Figure 2.13).

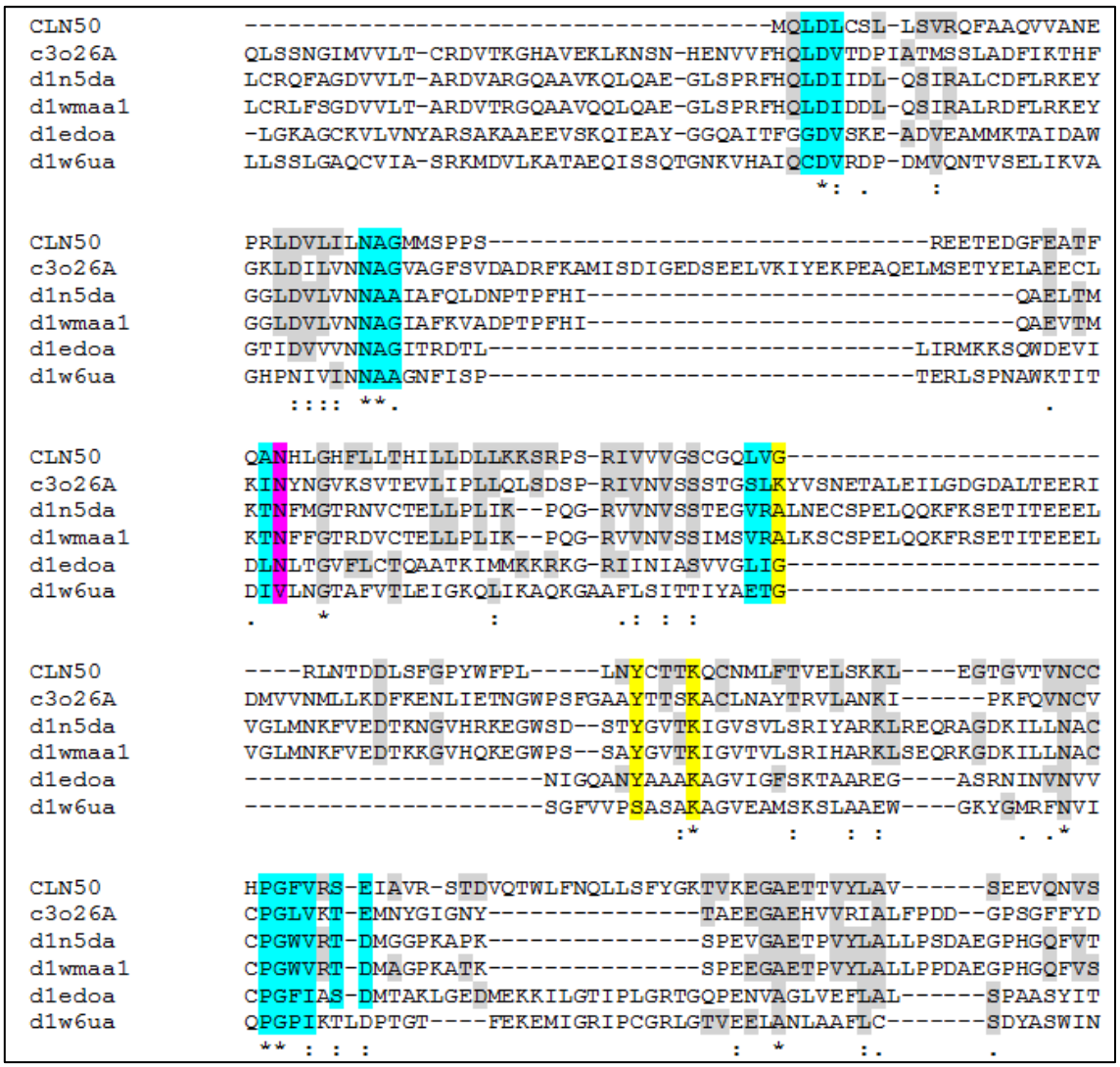


Figure 2.13: Section of amino acid alignment of putative tick RDH (clone 50) with threading hits. *Light grey represents identical nucleotides, blue represents NAD binding sites and yellow represent active residues. The YXXXK active site motif is underlined with red. Purple represents a possible cofactor binding or a catalytic residue. c3o26A: Oxidoreductase, Papaver somniferum (opium poppy), 99.9% confidence, 25% identity; d1n5da: Oxidoreductase, Sus scrofa (pig), 99.9% confidence, 21% identity; d1wmaa1: Oxidoreductase, Homo sapiens (human), 99.9% confidence, 24% identity; d1edoa: Oxidoreductase, Brassica napus (rapeseed), 99.9% confidence, 16% identity; d1w6ua: Oxidoreductase, Homo sapiens (human),99.9% confidence, 14% identity.*

Using Phyre[2], secondary structure prediction was performed and a structure rich in alpha-helical motifs was indicated. The low amount of disordered regions (0.1%) shows that the structure analysis was successful for finding the possible folds for most of the amino acid sequence (Figure 2.14) (Dunker *et al.*, 2001).

Figure 2.14: Secondary structure report of putative tick RDH clone 50 from Phyre[2]. *Shown is the amino acid sequence of clone 50's insert, its predicted secondary structure, predicted disordered regions and the confidence. The confidence of the predictions is color-coded from high to low (red to blue).*

Threading analysis was able to predict the tertiary structure of the protein of interest with confidence despite the low identity (30%) with the template sequence. The predicted tertiary structure of the putative *R. microplus* RDH is shown in Figure 2.15. Validation of predicted structure will be performed in future analyses.

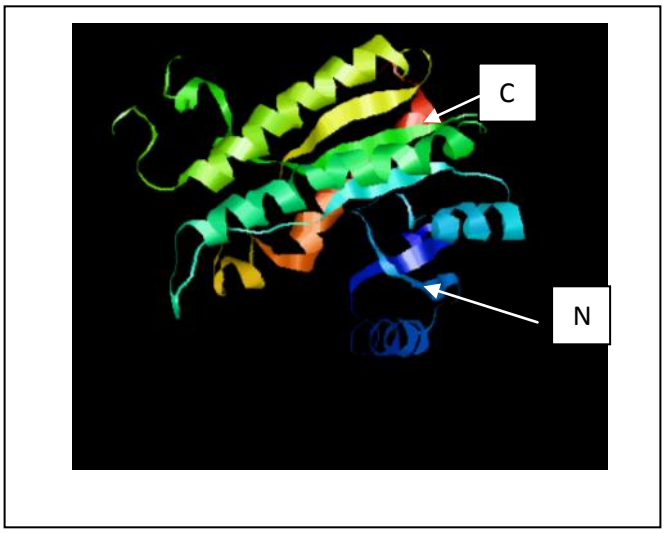


Figure 2.15: Predicted 3D structure of putative tick RDH clone 50 by Phyre[2]. *The 5' amino terminus (N) in blue and the 3' carboxyl terminus (C) in red is indicated. These are indicated with N and C respectively. The rest of the amino acid residues (5' to 3') are indicated with violet / blue / green / yellow / orange / red.*

BmATAQ Clone 12, putative Aldehyde dehydrogenase like protein (ALDH)

The amino acid alignment of clone 12 with the best hits from the PSI-BLAST showed very weak sequence identity. However, the best hit representing aldehyde dehydrogenase family 16 member A1 had an expected value of 5*E*-13 value. Clone 12, putative tick aldehyde dehydrogenase was found to be closest to the *Bos Taurus* (cattle) homolog, ALDH16A1 (Figure 2.16).

The active site of an ALDH has the following motif: [Leu Ile Val Met Phe Gly Ala] – **Glu** – [Leu Ile Met Ser Thr Ala Cys] – [GlySer] – **Gly** – [Lys Asn Leu Met] (Maeda *et al.*, 2000). To allow binding with coenzymes like NADH, oxidoreductases have the following motif: (Gly or Ala)-(Xaa)3-(Gly, Ala or Thr)-(Ile or Leu)-(Gly, Ala or Ser)-(Xaa)10-(Gly or Asn) (Liu *et al.*, 1997). Some of these functional sites could be identified on clone 12 with an alignment analysis using the best hits obtained from PSI-BLAST (Figure 2.16)

Figure 2.16: Amino acid alignment of putative *R. microplus* ALDH (clone 12) with PSI-BLAST hits. *Nucleotides highlighted with yellow indicates NAD+ binding sites, with blue indicating substrate binding sites and nucleotides highlighted in green indicating the catalytic site. GI:166198271; ALDH16A1, Bos taurus (cattle), E value of 5E-13; GI:269849532, ALDH16A1, Homo sapiens (human), E value of 2E-09; GI:166198365, ALDH16A1, Mus musculus (house mouse), E value of 1E-11; GI: 123780915: ALDH16A1, Rattus norvegicus (Norway rat), E value of 4E-11; GI: 2506347: Aldehyde dehydrogenase, Escherichia coli (bacteria), E value of 2E-06.*

Blast analysis in tick data base (BmGI) supported the result from PSI-BLAST analysis. However, sequences had very weak identity as seen from PSI-BLAST (Appendix 2). The best hit from BmGI was TC16791 with 60% similarity (E value of E0.04) and was used for further sequence analysis to prove existence of the catalytic and cofactor binding sites in putative aldehyde dehydrogenase (Figure 2.17).

Figure 2.17: Amino acid alignment of ALDHs from the BMGI best hit (TC16791) with its PSI-BLAST hits. *Grey; identical nucleotides. Upstream cofactor NAD+ binding site (IIPW) is written in red. Blue; glutamic acid active site, the active residue E is indicated in red, yellow; cysteine active site, the active residue C is indicated in green. GI: 3121992: Mitochondrial aldehyde dehydrogenase, Mesocricetus auratus (golden hamster), E value of E-00; GI: 118505: Mitochondrial aldehyde dehydrogenase, Rattus norvegicus (Norway rat), E value of E-00; GI: 118504: Mitochondrial aldehyde dehydrogenase, Homo sapiens (human), E value of E-00.*

Threading analysis showed that clone 12 is possibly an oxidoreductase related to an ALDH-like protein (c3hazA), from *Bradyrhizobium japonicum*. *B. japonicum* is a legume-root nodulating microsymbiotic nitrogen-fixing bacterium (Bonnet *et al.*, 2013). Though, threading analysis gave a result with 100% confidence and 75% coverage, only 17% sequence identity is shared between these two proteins (Figure 2.18).

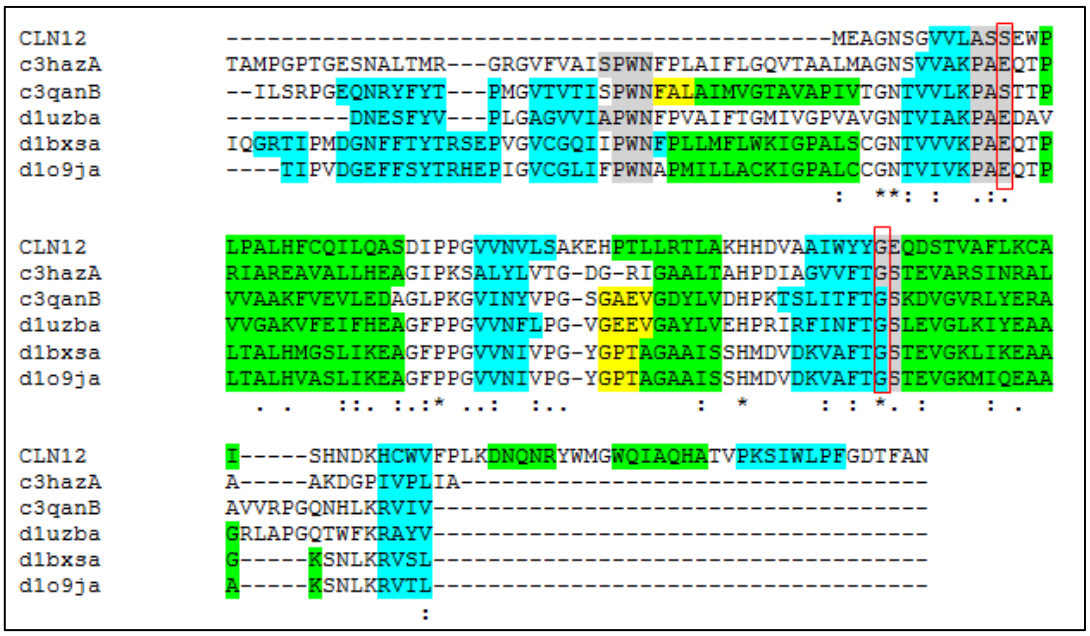


Figure 2.18: Amino acid alignment of putative tick ALDH (Clone 12) with threading hits. *Blue: beta strands, green: helix turns, yellow: turns, and grey: possible binding sites or active residues, two of the possible active sites are indicated with the red blocks. c3hazA: Oxidoreductase, ALDH-like protein, Bradyrhizobium japonicum USDA 110, 100% confidence, 17% identity;c3qanB: Oxidoreductase, ALDH-like protein, Bacillus halodurans (bacteria), 100% confidence, 19% identity; d1uzba: Oxidoreductase, ALDH-like protein, Thermus thermophilus HB8, 100% confidence, 19% identity; d1bxsa:Oxidoreductase, ALDH-like protein, Ovis aries (sheep), 100% confidence, 23% identity; d1o9ja:Oxidoreductase, ALDH-like protein, Elephantulus edwardii (Cape elephant shrew) 100% confidence, 22% identity.*

Similarity between the location of secondary structures such as beta sheet and alpha helixes of the best hits of threading analysis indicates that clone 12 could possibly fold into a similar tertiary. Correlation between putative active residues in clone 12 also supports this hypothesis to a certain degree (Figure 2.18).

Secondary structure prediction of clone 12, using Phyre[2], indicated a structure rich in both alpha-helices and β-sheets. Disordered regions are less abundant indicating that the tertiary structure for the majority of the sequence could be predicted with a high degree of confidence (Figure 2.19).

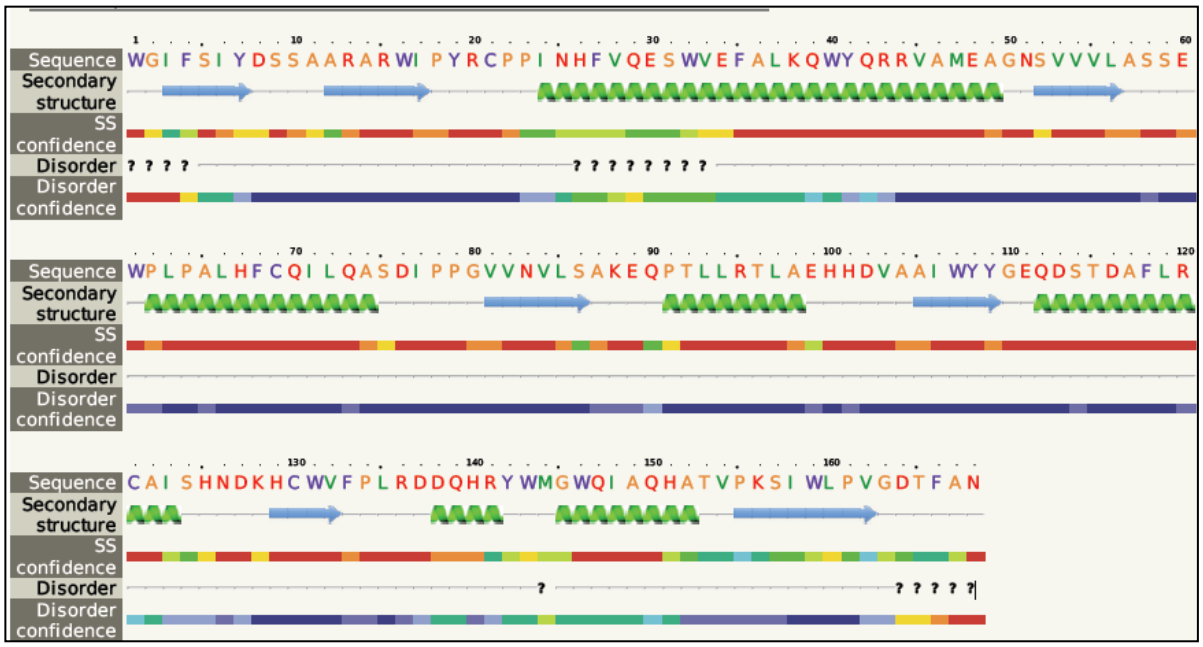


Figure 2.19: Secondary structure report of clone 12 from Phyre[2]. *The secondary structure for clone 12's insert was predicted with high confidence and coverage. The predicted structure is rich in both alpha-helices and β-sheets with very little disorder (unpredictable secondary structure).*

The tertiary structure of clone 12 is shown in Figure 2.20, as predicted using Phyre. Using structural analysis, it was possible to predict a secondary structure with confidence for the entire protein, despite the low sequence identity between the query and template sequences (<20%). This is possible due to proteins' being more conserved on structural level compared to sequence level.

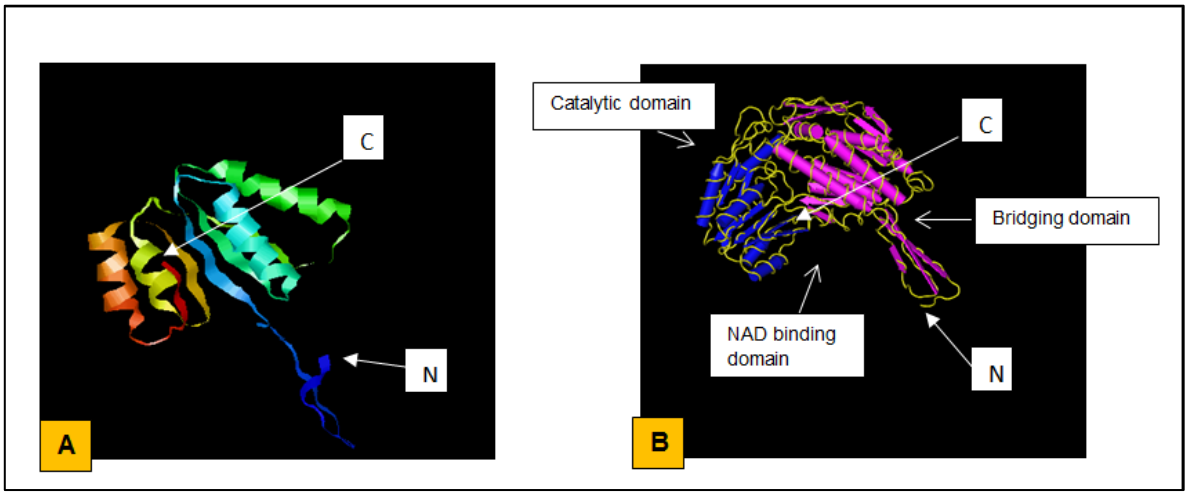


Figure 2.20: Predicted 3D structure of clone 12 (ALDH) by Phyre[2](A) and modified 3D structure of a typical ALDH (B) (Liu *et al.*, 1997). *The amino terminus (N) is indicated in blue and the carboxyl terminus (C) is indicated in red. A ribbon drawing of an ALDH monomer with its domains and secondary structure is shown in figure B.*

Aldehyde dehydrogenases catalyze the pyridine nucleotide-dependent oxidation of aldehydes to acids. They are highly reactive molecules that are intermediates or products involved in a broad spectrum of physiological and biological processes. These enzymes are NAD(P)-dependent and are commonly found in almost all animals and plants (Vasiliou *et al.*, 2000). To date, 5 archeal, 26 eubacterial and 17 eukaryotic ALDH genes have been reported, and 22 eukaryotic gene families have been identified (Yang *et al.*, 2012). An active molecule of ALDH is a dimer containing two identical subunits. Each subunit in turn consists of a catalytic domain, (arm-like) bridging domain and an NAD(P)-binding domain (Figure 2.20). The polypeptide chain loops to start the bridging domain and returns to form the Rossmann fold, open α/β dinucleotide-binding structure. The catalytic domain also has an α/β pattern with

six parallel and one anti-parallel β-sheet and five α-helices. The chain finally develops back over the NAD-binding domain (Figure 2.20) (Alberts *et al*., 2002).

Bm86 Clone 93, putative Aldehyde dehydrogenase like protein (ALDH)

PSI-BLAST analysis revealed that clone 93 had homology with the aldehyde dehydrogenase family 16 member A1 (ALDH16A1) from *Bos taurus* with an E value of 5E-13.Sequence similarities in aligned amino acid sequences were low as it was 30% with the PSI-BLAST analysis' best hit. Amino acid alignment analysis showed that clone 93, a putative *R. microplus* ALDH-like protein, has some of the characteristics NAD- binding and active site residues (Figure 2.21). However, the consensus motif of the active site of ALDH, [Leu-Ile-Val-Met-Phe-Gly-Ala] – **Glu** – [Leu-Ile-Met-Ser-Thr-Ala-Cys] – [Gly-Ser] – **Gly** – [Ly- Asn-Leu-Met] could not be identified (Maeda *et al.*, 2000). The most common amino acid sequence necessary for the capability of utilizing NADH as a coenzyme is (Gly or Ala)-(Xaa)3-(Gly, Ala or Thr)-(Ile or Leu)-(Gly, Ala or Ser)-(Xaa)10-(Gly or Asn) (Liu *et al.*, 1997).

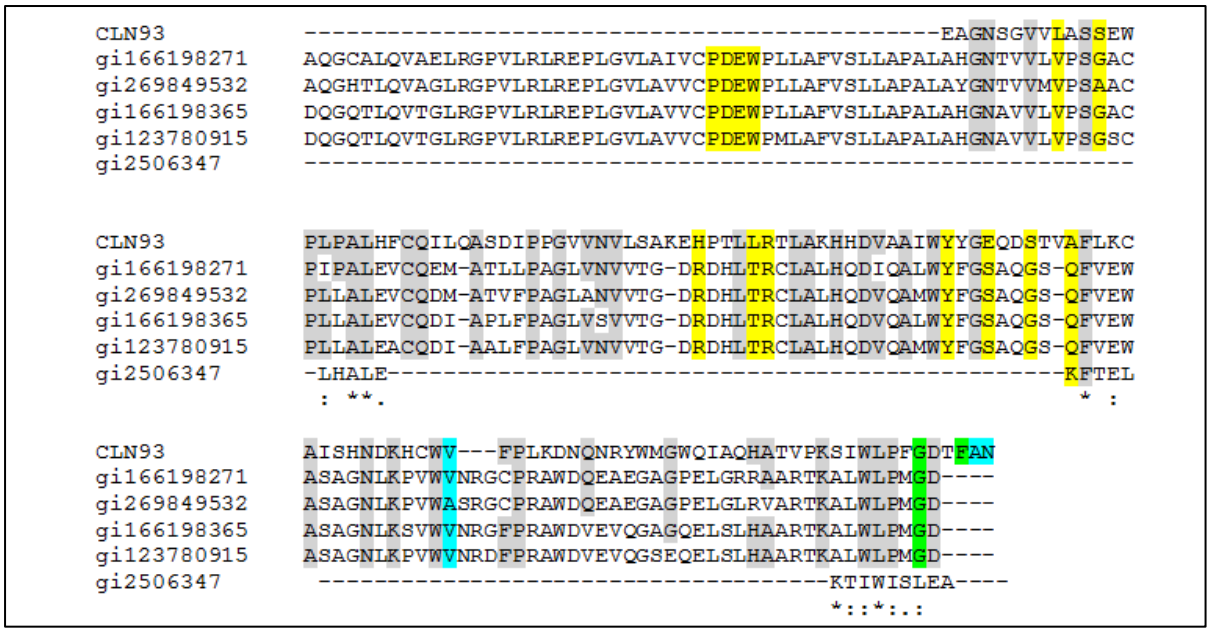


Figure 2.21: Amino acid alignment of putative *R. microplus* ALDH-like protein (CLN93) with PSI-BLAST hits. *Yellow nucleotides indicate NAD+ binding sites, blue nucleotides indicate substrate binding sites and green nucleotides indicate catalytic sites.GI: 166198271; ALDH16A1, Bos taurus (cattle), E value of 5E-13; GI: 269849532, ALDH16A1, Homo sapiens (human), E value of 2E-09; GI: 166198365, ALDH16A1, Mus musculus (house mouse), E value of 1E-11; GI: 123780915: ALDH16A1, Rattus norvegicus (Norway rat), E value of 4E-11; GI: 2506347: Aldehyde dehydrogenase, Escherichia coli (bacteria), E value of 2E-06.*

A BLAST search against the BmGI database also revealed that clone 93 is a putative ALDH-like protein. The sequence identity between the BmGI best hit (TC16791) and clone 93 was found to be 60%. TC16791 shares sequence similarity to an aldehyde dehydrogenase (class 2) family from *Xenopus tropicalis* (Western clawed frog). Both the aldehyde dehydrogenase glutamic acid catalytic residue and the aldehyde dehydrogenase cysteine catalytic residue are found in TC16791 (Figure 2.22).

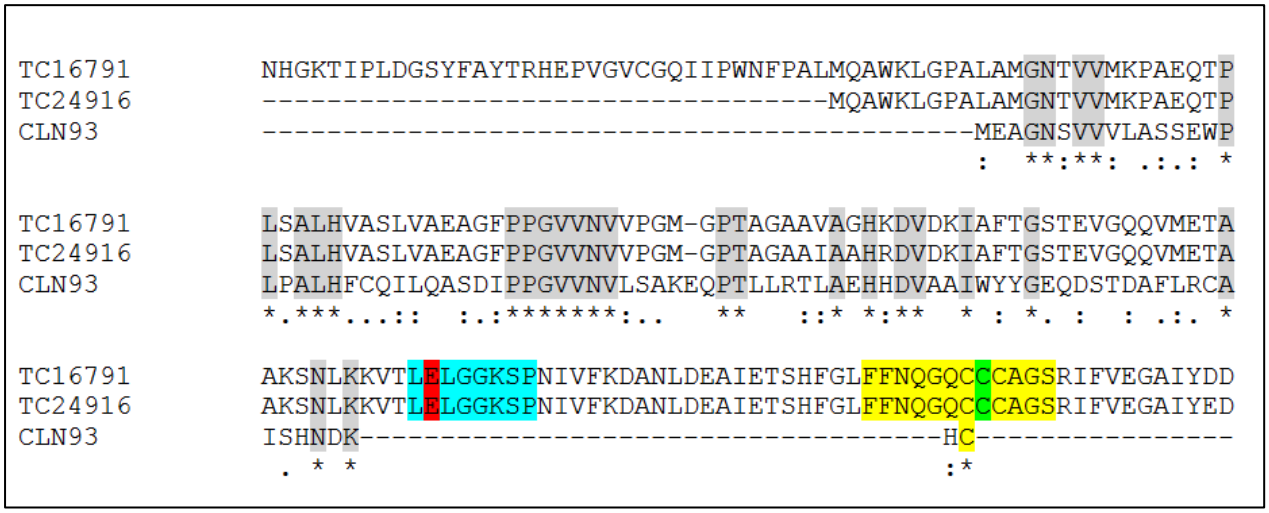


Figure 2.22: Amino acid alignment of putative *R. microplus* ALDH-like protein (CLN93) and BMGI best hits. *Aldehyde dehydrogenase glutamic acid active site is indicated in blue and the active residue is indicated in red. The aldehyde dehydrogenase cysteine active site is indicated in yellow and the active residue is indicated in green. Identical amino acids between aligned sequences are highlighted with grey.TC16791: Aldehyde dehydrogenase 2 family; Xenopus tropicalis (Western clawed frog), E value of E0.00012; TC24916: Mitochondrial aldehyde dehydrogenase, Rattus norvegicus (Rat), E value of E0.0004.*

To identify putative NAD+ binding sites the TC16791 amino acid sequence was further analyzed with PSI-BLAST (Figure 2.23).

Figure 2.23: Amino acid alignment of putative tick ALDH (TC16791) with PSI-BLAST hits. *Grey; identical nucleotides. Upstream cofactor NAD+ binding site (IIPW) is written in red. Blue; glutamic acid active site, the active residue E is indicated in red, yellow; cysteine active site, the active residue C is indicated in green. GI: 3121992: Mitochondrial aldehyde dehydrogenase, Mesocricetus auratus (golden hamster), E value of E-00; GI: 118505: Mitochondrial aldehyde dehydrogenase, Rattus norvegicus (Norway rat), E value of E-00; GI: 118504: Mitochondrial aldehyde dehydrogenase, Homo sapiens (human), E value of E-00.*

Structural prediction using Phyre[2] showed that clone 93 is an oxidoreductase, with similarity to, an ALDH-like protein (c3hazA) from *Bradyrhizobium japonicum*, with a 100% confidence due to a conserved secondary structure (Figure 2.24). Furthermore, secondary structure prediction of this putative tick ALDH, indicated a structure rich in both alpha-helices and β-sheets. Two of the possible active site residues were observed, as well as some of the predicted binding site residues though sequence identity between the aligned sequences were very weak at 16 % (Figure 2.24). This is another example of a strong conservation of structure with low sequence identity.

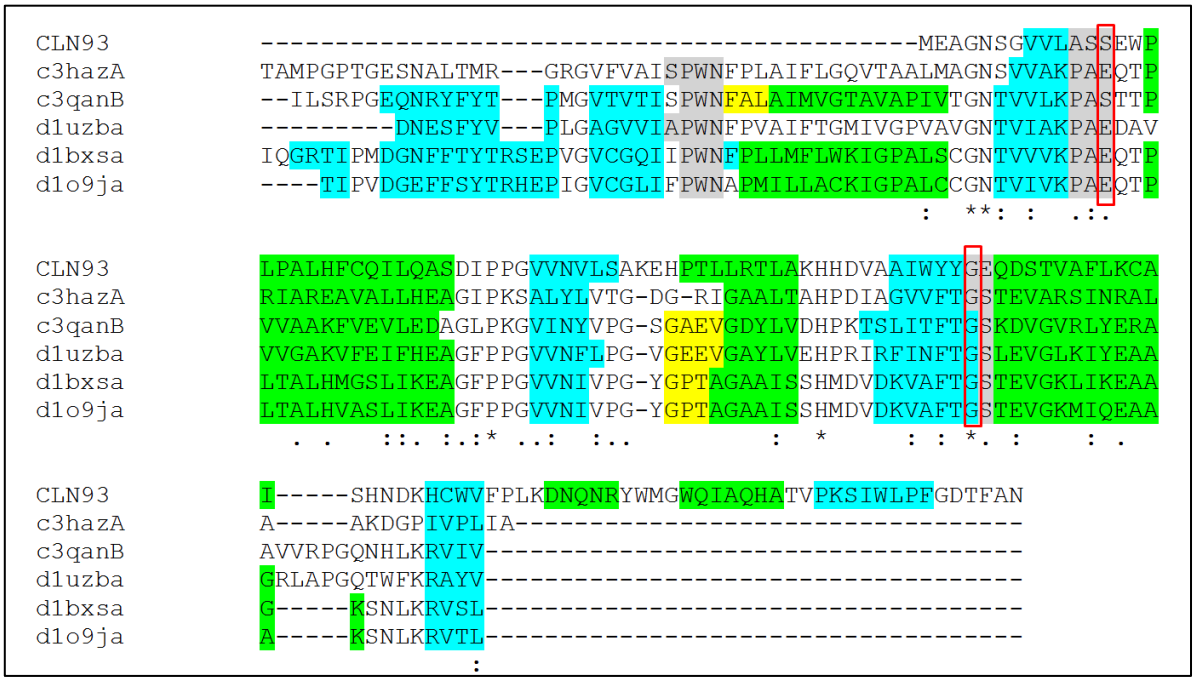


Figure 2.24: Amino acid alignment of putative *R. microplus* ALDH-like protein (CLN93) with threading hits. *Blue: beta strands, green: helix turns, yellow: turns, and grey: possible binding sites or active residues, two of the possible active sites are indicated with the red blocks. c3hazA: Oxidoreductase, ALDH-like protein, Bradyrhizobium japonicum USDA 110, 100% confidence, 17% identity; c3qanB: Oxidoreductase, ALDH-like protein, Bacillus halodurans (bacteria), 100% confidence, 19% identity; d1uzba: Oxidoreductase, ALDH-like protein, Thermus thermophilus HB8, 100% confidence, 19% identity; d1bxsa:Oxidoreductase, ALDH-like protein, Ovis aries (sheep), 100% confidence, 23% identity; d1o9ja:Oxidoreductase, ALDH-like protein, Elephantulus edwardii (Cape elephant shrew) 100% confidence, 22% identity.*

Sequence analysis of clone 12 (putative BmATAQ interacting partner) and clone 93 (putative Bm86 interacting partner) has shown identical results. It can therefore be hypothesized that clone 12 and clone 93 are the same, ALDH-like protein.

Bm86 Clone 53, putative multi-domain Kunitz-like protein

Following PSI-BLAST analysis, clone 53 was found to encode a multi-domain Kunitz-like protein. The 12 conserved cysteine residues were identified that suggest the presence of two possible Kunitz domains (Figure 2.25).

Figure 2.25: Amino acid alignment of predicted *R. microplus* Kunitz-like protease inhibitor (CLN53) with PSI-BLAST hits. *The cysteine residues of two Kunitz domains are highlighted in green; trypsin interaction site is indicated in yellow which include the second cysteine residue of first Kunitz domain. GI: 82203423: Disulfide bond, Serine protease inhibitor, Bitis gabonica (Gaboon viper),5e-21; GI: 114152125: Serine protease inhibitor, Drosophila melanogaster (fruit fly), E value of 2E-18; GI:74935652: Serine protease inhibitor, Rhipicephalus microplus (cattle tick), E value of 8E-19; GI: 2497697: Serine protease inhibitor, Rattus norvegicus (Norway rat), E value of 9E-19; GI:172046756: Serine protease inhibitor, Mus musculus (house mouse), E value of 9E-19.*

BLAST searches of the deduced amino acid sequence of clone 53 against the tick database and the *R. microplus* Gene Index (BmiGi), confirmed it is a possible multi Kunitz domain protein. Clone 53 had a match with CK176583, which was predicted to be similar to Kunitz-type proteinase inhibitor 1 (SHPI-1) from *Stichodactyla helianthus* (Carribean sea anemone). However, the first cysteine residue of the (N terminal) first Kunitz-domain of clone 53 was not identified in the sequence due to sequence truncation from cDNA synthesis. Although the sequence is truncated, it is valuable as it shows the region binding to the bait.

The BPTI-Kunitz domain spans approximately 60 amino acids containing 6 cysteine residues (XnCysX8CysX15CysX7CysX12CysX3CysXn) forming a specific pattern of 3 disulfide bridges between Cys1-Cys6, Cys2-Cys4 and Cys3-Cys5 (Figure 2.25) (Francischetti *et al.*, 2005). It comprises a single 3_{10}-helix around the N-terminal cysteine with the central part of the fold forming a double-stranded anti-parallel β-sheet. The anti-parallel strands of the β-sheet are linked by a β- hairpin loop and the structure ends with a C-terminal three turn α-helix (Ascenzi *et al.*, 2003). These Kunitz domain-containing proteins mostly target trypsin-like serine proteases such as those in the blood coagulation cascade, although mutational studies have shown that their protease specificity can be altered (Francischetti *et al.*, 2004).

Trypsin-like serine protease inhibitors have a functional arginine (R) in the P1 position (Macedo-Ribeiro *et al.*, 2008b). Therefore, for clone 53 it is possible that the P1 residue is an arginine in the first Kunitz domain and glutamine in the second Kunitz-domain, both located after the second cysteine residue of each domain (Figure 2.26). Binding of Kunitz-type inhibitors to thrombin is enhanced by Glu substitution at the P1 active site (Francischetti *et al.*, 2004; Macedo-Ribeiro *et al.*, 2008a; Paesen *et al.*, 2009).

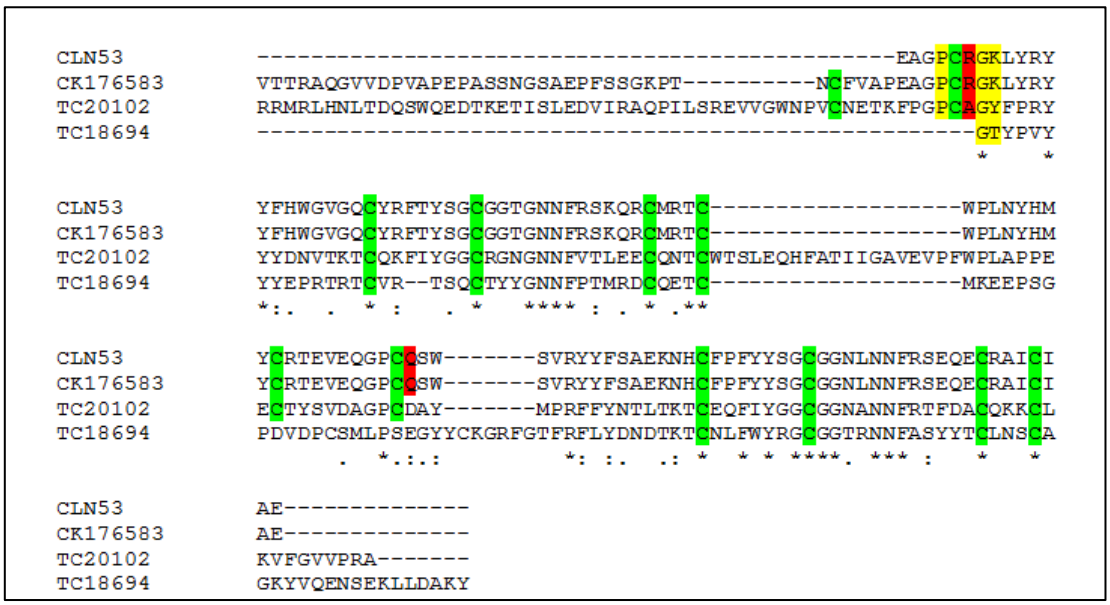


Figure 2.26: Amino acid alignment of putative *R. microplus* Kunitz-like protease inhibitor (CLN53) with BMGI hits. *The cysteine residues of the first Kunitz domain 1 and 2 are highlighted with green and yellow, respectively. The predicted P1 active sites are highlighted with red for both domains. CK176583: Kunitz-type proteinase inhibitor SHPI-1, Stichodactyla helianthus (Carribean sea anemone), E value of 3E-69; TC20102: Kunitz-type serine protease inhibitor, Rhipicephalus microplus (Cattle tick), E value of E=0.51, TC18694: Papilin proteoglycan-like sulfated glycoprotein; Mus musculus (house mouse), E value of E=0.75.*

Threading analysis of clone 53 using Phyre[2]also showed that this putative Bm86 binding partner has a Kunitz-type folding structure. The best hit with the highest sequence identity (39%) was c2odyF with 100% confidence (Figure 2.27). C2odyF is chain F of a functional and accessible reactive-site loop found in thrombin-bound Boophilin from *Rhipicephalus microplus* (cattle tick). An additional hit from a soft tick species (c1tocR) was identified that encodes a serine Proteinase from *Ornithodoros moubata* (soft tick, eyeless tampan) with 100% confidence and 16% amino acid sequence identity.

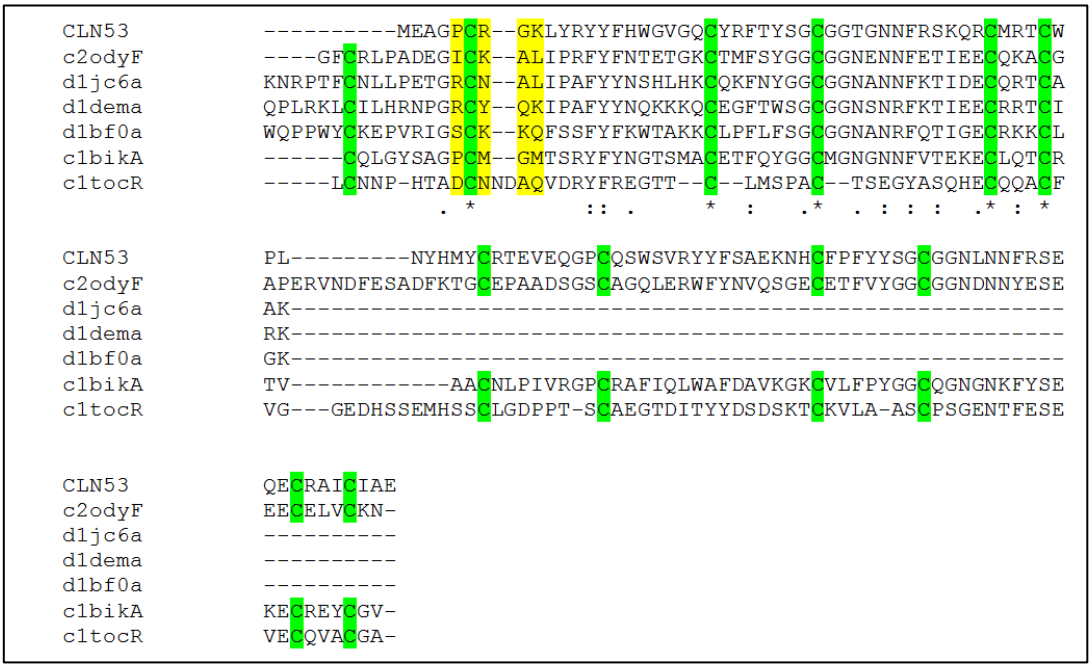


Figure 2.27: Amino acid alignment of putative *R. microplus* Kunitz-like protease inhibitor (CLN53) with threading hits. *The cysteine residues of the first Kunitz domain 1 and 2 are highlighted with green. The trypsin interaction site indicated with yellow. c2odyF, GI :166007042: Thrombin-bound Boophilin, Rhipicephalus microplus (cattle tick), 100% confidence, 39% sequence identity; d1jc6A, GI: 33357119:Kunitz-Type Chymotrypsin Inhibitor, Bungarus fasciatus (banded krait), 99.8% confidence, 35% identity; d1dema, GI: 157830809: Proteinase Inhibitor, Dendroaspis polylepispolylepis (black mamba), 99,8% confidence, 33% identity; d1bf0a, GI: 157830265: BPTI/Kunitz family of serine protease inhibitors, Dendroaspis angusticeps (eastern green mamba), 99,8 confidence, 37% identity; c1bikA, GI: 4699843: serine protease inhibitor with two Kunitz domains Homo sapiens (human), 100% confidence, 33% identity; c1tocR, GI: 2392576: serine proteinase, Ornithodoros moubata (tick), 100% confidence, 16% identity.*

The predicted three dimensional structure of *R. microplus* trypsin-like serine protease inhibitor has two prothrombin proteins which interact with prothrombin, phosphate ion, sodium ion and boopilin (www.ncbi.nlm.nih/). The structure also has two Kunitz domains which interact with prothrombin (Figure 2.28).

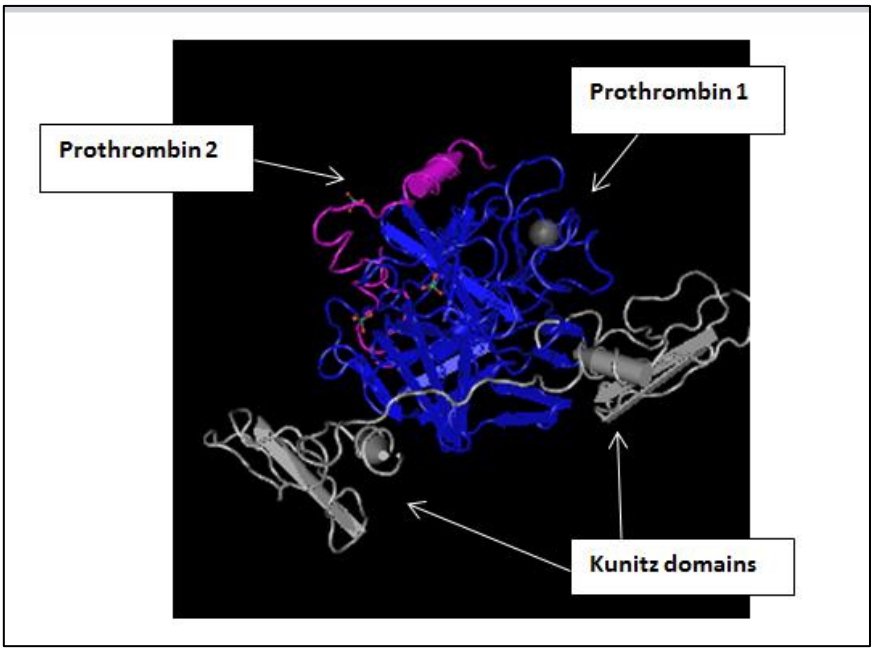


Figure 2.28: Molecular representation of the three dimensional structure of putative *Rhipicephalus microplus* Kunitz-like protease inhibitor. *Prothrombin 1 belongs to thrombin light superfamily which interacts with phosphate ion and prothrombin. Protrombin 2 belongs to Trypsin-like serine protease (trp-SPc) superfamily which interacts with boohilin, phosphate ion, prothrombin and sodium ion. Kunitz domains belong to boophilin and interact with prothrombin. (Figure taken from: www.ncbi.nlm.nih/).*

2.4.5. Immunoprecipitation of BmATAQ and Bm86 interacting candidates

The binding partners of BmATAQ and Bm86 were isolated from tick gut proteins using immunoprecipitation with agarose-bound protein A, after which they were analyzed using SDS-PAGE or stored at -20°C.

BmATAQ binding partners

The interacting partners isolated by immunoprecipitation were analyzed using SDS-PAGE and subsequently stained using Coomassie blue. The presence of the putative binding partners was detected and confirmed by comparison of the corresponding molecular weight of the proteins visualized on the gel. The average molecular weight of ALDH16A1, RDH14 and BmATAQ are known to be 85 kDa, 32-39 kDa and 66 kDa, respectively, corresponding with results from SDS-PAGE. The presence of RDH14, as a BmATAQ interacting partner, was also confirmed with Western blot analysis (Figure 2.29 E).

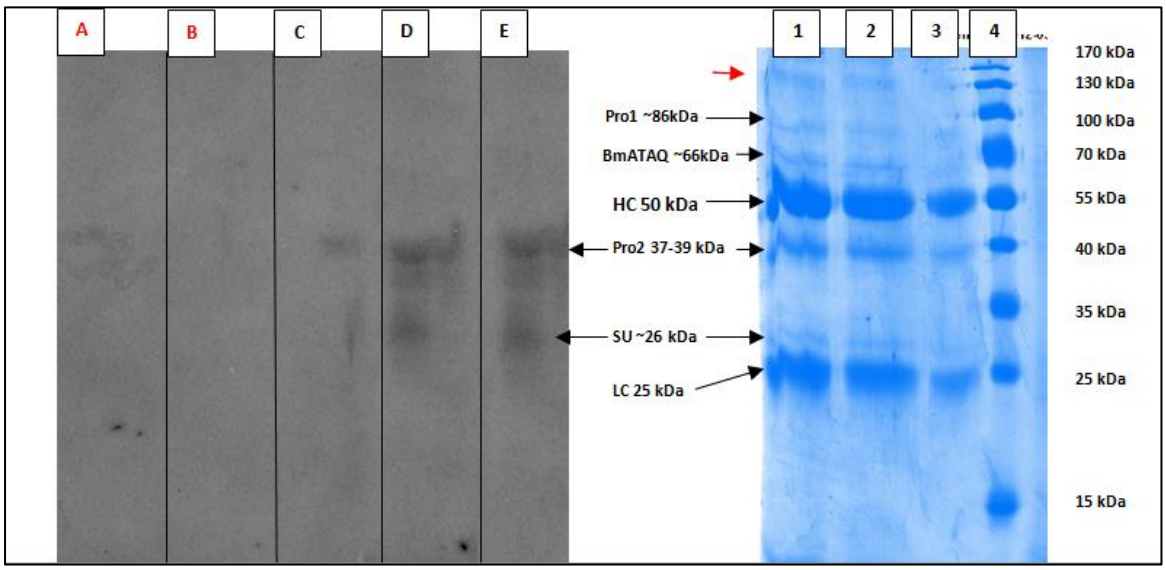


Figure 2.29: Comparison of the chemiluminescent film which was exposed to a PVDF membrane transferred with BmATAQ and its binding partners (A), with the Coomassie blue stained polyacrylamide gel of BmATAQ binding partners (B). *Lane A and B are repeat of negative control for unspecific binding of secondary antibody; lane C-E are different time point of exposure, 30 minutes, 10 minutes and 3 minutes, respectively; lane 1-3 are different dilutions of BmATAQ and its binding partners; lane 4is a PageRuler™ Prestained Protein Ladder. The red arrow represent either be a novel BmATAQ-binding partner or bovine IgG (heavy and light chains). The band with size 86 kDa (Prot 1) could be predicted as R. microplus ALDH-like protein. Bands with 66, 50 and 25 kDa weight are predicted as BmATAQ, heavy chain (HC) and light chain (LC) of IgG, respectively. Bands with 37-39 kDa and 26 kDa weights are predicted as R. microplus RDH and the RHD subunit, respectively.*

Following optimization of the Western blot, two protein bands could be detected. Results indicate that RDH is present in two forms: a native form which is 37-39 kDa and as a subunit 26 kDa in size. Both forms are visible in Figure 2.29 after both SDS-PAGE and Western blot analysis. The polyacrylamide gel indicated the presence of one extra BmATAQ binding partner that could not be obtained from previous Yeast two-hybrid screening (Figure 2.29). This protein could also represent IgG (heavy and light chains), two heavy chains (55x2=110 kDa) and two light chains (25x2=50 kDa) that together could result in a 160 kDa protein (Figure 2.29).

Bm86 interacting partners

SDS-PAGE confirmed the presence of the putative interacting partners of Bm86, based on their molecular weight. The predicted molecular weight of putative *R. microplus* Kunitz-like protein is 22 kDa, while the molecular weights of both putative *R. microplus* ALDH-like protein and Bm86 are known to be 86 kDa. Proteins corresponding to the heavy and light chains of bovine IgG are indicated on figure 2.30. Putative *R. microplus* Kunitz-like protein (Prot 1) is only visible in the membrane fraction and has a molecular weight very close to that of the light chain of bovine IgG. The reason why clone 53 is not seen in the soluble fraction might be due to the high amounts of the light chain as they have same molecular weight.

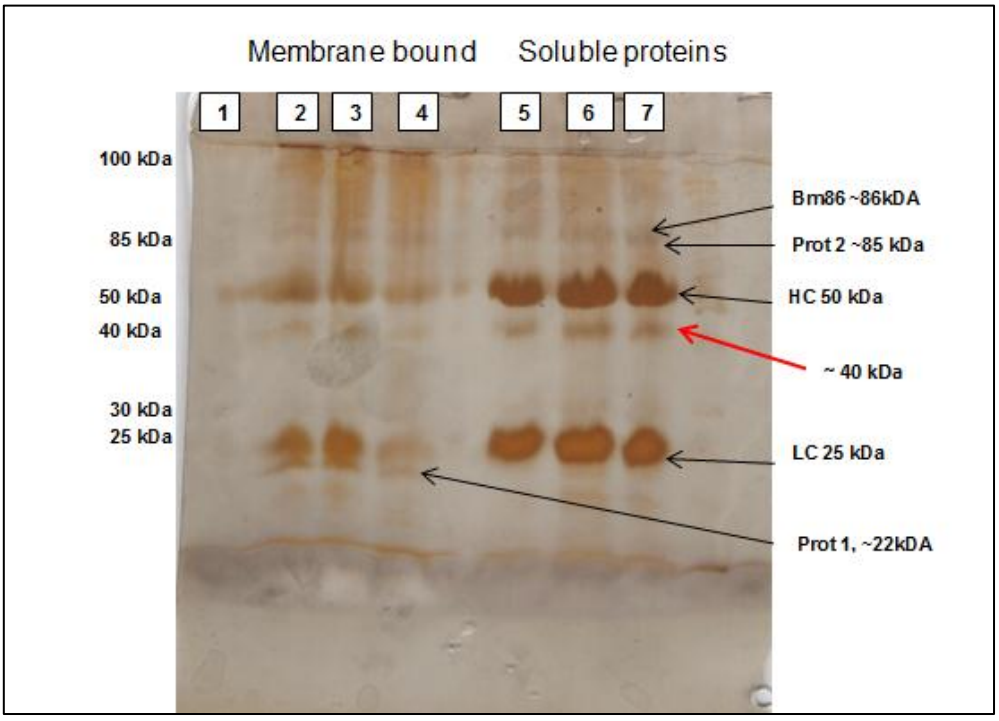


Figure 2.30: Silver-stained polyacrylamide gel of immunoprecipitated Bm86 and its interacting partners. *Lane 1= unstained protein ladder; lanes 2-4 = membrane-bound Bm86 interacting proteins in different dilutions; lane 5-7= soluble Bm86 interacting proteins in different dilutions. Prot 1= Putative R. microplus Kunitz-like protein (clone 53); prot 2 = putative R. microplus ALDH-like protein (clone 93). The protein indicated with red arrow (~ 40 kDa) is a putative unknown interacting partner of Bm86.*

Bm86 interacting partners, isolated during the second round of immunoprecipitation, were separated using SDS-PAGE and stained with Coomassie blue, which is compatible with downstream MS analysis. When Coomassie-stained polyacrylamide gels were examined, clone 53's product could not be separated from the IgG light chain as they have similar molecular weight. Considering the lower abundance of clone 53's product compared to the light chain of bovine IgG, the protein samples were separated using tricine gels, which are capable of separating proteins from 1-100 kDa. Clone 53's product, however, could not be obtained using tricine-based SDS-PAGE either. This could be because of a weaker interaction between Bm86 and clone 53's product (Kunitz-like protein), which was lost during the washing steps or this protein is only found among the membrane-bound proteins from gut. This has to be repeated in further studies on pure Bm86 interacting partners. Considering the size similarity between BmATAQ and Bm86 interacting partners it is hypothesized that RDH might be an interacting partner for both Bm86 and BmATAQ (Figure 2.31).

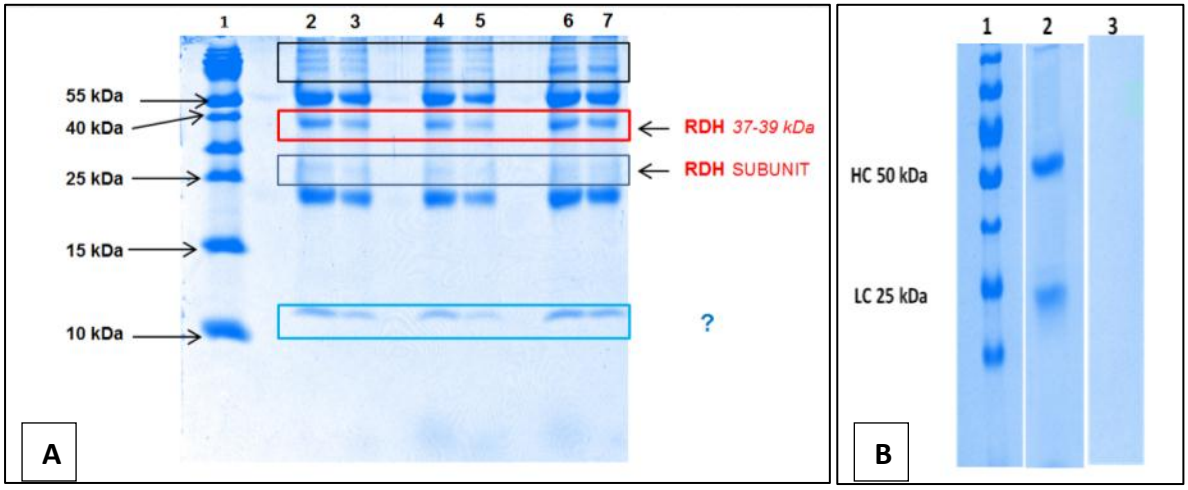


Figure 2.31: Tricine gel of Bm86 membrane-bound (lane 2 and 3), Bm86 soluble (lane 3 and 4) and BmATAQ (lane 6 and 7) interacting proteins (A). *The possible common interacting partners of BmATAQ and Bm86 are indicated and the unknown binding partner is shown with a question mark. RDH seems to be a common binding partner, which has to be confirmed with Western blot. Lane1is prestained protein ladder.* **The controls are shown in picture (B).** *Lane 1; prestained protein ladder, Lane 2; boiled serum sample, no tick gut protein, Lane 3; back ground, no sample.*

Western blot analysis was done on both membrane bound and soluble gut protein against RDH14 antibody to detect the presence of RDH14 in the Bm86 binding proteins (Figure 2.32). The subunit of RDH14 was obtained following Western blot analysis of Bm86 interacting partners as well.

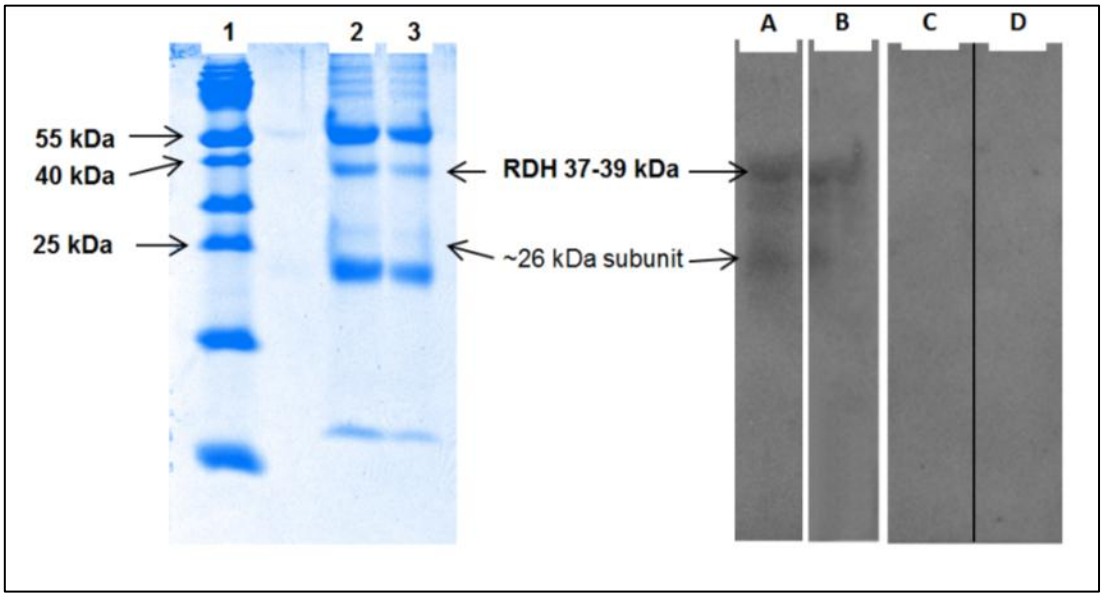


Figure 2.32: Comparison of tricine SDS-PAGE (left) and Western blot analysis of Bm86 interacting partners using RDH antibodies (right). *Lane1; prestained protein ladder, Lanes 2 and 3; Bm86 membrane-bound and soluble interacting proteins respectively. LanesA and B; respective western blot results; Lanes C and D; respective negative controls.*

LC-MS/MS

Interacting partners of BmATAQ and Bm86, separated using SDS-PAGE, were analyzed with LC-MS/MS to confirm the interacting partners and to obtain their amino acid sequences. The presence of BmATAQ and Bm86 was confirmed with mass spectrometry analysis. However, it was problematic to obtain proof of the interacting partners. Although the amino acid sequences of the interacting partners could not be determined, MS analysis did indicate the presence of oxidoreductases for both BmATAQ and Bm86 binding proteins (Table 2.4).

Table 2.4: Selected LC-MS/MS results showing presence of dehydrogenase and oxidoreductase. P00562; GI:416596; Dehydrogenase from *Escherichia coli* (bacteria),**O57672**; GI:6016081;Oxidoreductase from *Meleagris gallopavo* (turkey),**P00349**; GI:112844;Oxidoreductase from *Ovis aries* (sheep).

Accession	Description	Organism	Number of peptides	Molecular weight kDa
P00562	Dehydrogenase	*Escherichia coli* (bacteria)	4	88
O57672	Oxidoreductase	*Meleagris gallopavo*(turkey)	1	25
P00349	Oxidoreductase	*Ovis aries*(sheep)	2	53

This provides tentative proof of the presence of both ALDH and RDH as interacting partners of both BmATAQ and Bm86as seen from western blot analysis. In a future study, the inserts can be cloned again and the interactions can be identified/confirmed with the yeast two-hybrid assay and confirmed with co-localization. Size-exclusion can also be used on purified protein and the retention of bound and unbound protein assessed.

Since BmATAQ and Bm86 have two common interacting partners, ALDH-like and RDH-like, it is most likely that these two target proteins are involved in the same metabolic pathway/s. The ALDH16 family encode enzymes commonly found in different types of tissues, such as bone marrow, kidneys, heart and lung (Marchitti *et al.*, 2008). These genes are found in protista, fish, bacteria and amphibians, however, remains to be determined in archaea, plants, and fungi (Vasiliou *et al.*,

2013). Further evaluation of the human ALDHs revealed a large variance, from < 15% to > 80%, in their amino acid identity (Persson *et al.*, 1991).

2.4.6. Dehydrogenase functions in retinoic acid metabolism

Both ALDH and RDH are involved in vitamin A metabolism (Figure 2.33). Vitamin A performs a role in embryonic development, vision, immune function, and tissue remodeling and metabolism (McCollum, Davis, 1913; Noy, 2006; Osborne, Mendel, 1919).

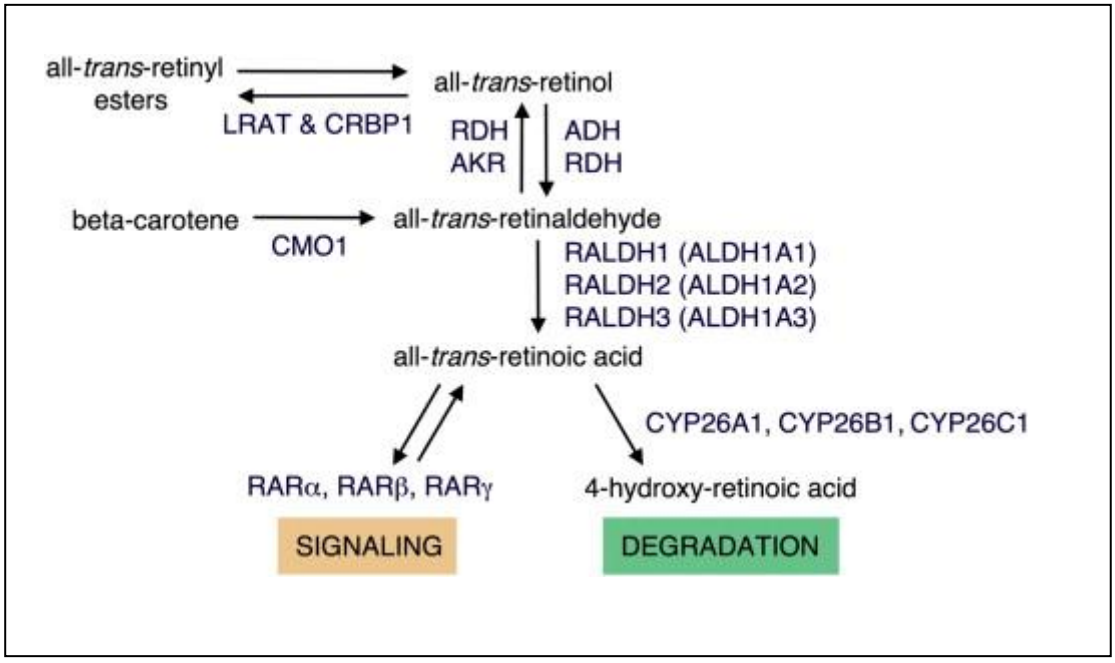


Figure 2.33: Retinoid metabolic pathway. *All-*trans*-retinol (the alcohol form of vitamin A) is converted to retinyl esters by lecithin-retinol acyltransferase (LRAT) and cellular retinol-binding protein-1 (CRBP1). All-*trans*-retinol is oxidized to all-*trans*-retinaldehyde by either alcohol dehydrogenase (ADH) or retinol dehydrogenase (RDH) using NAD as cofactor. All-*trans*-retinaldehyde is reduced to all-*trans*-retinol by RDHs using NADPH as cofactor or by aldo-ketoreductases (AKR). All-*trans*-retinaldehyde is further oxidized to all-*trans*-retinoic acid (RA) by retinaldehyde dehydrogenases (RALDH1, RALDH2, and RALDH3). RA is further oxidized to 4-hydroxy-retinoic acid by cytochrome P450s (CYP26A1, CYP26B1, and CYP26C1). RA initiates a signaling event via binding to nuclear RA receptors (RARα, RARβ, and RARγ) that regulate transcription of target genes (Figure taken from:www.sciencedirect.com [2012, October 2]).*

The parental vitamin A molecule is called retinoid. The principal circulating retinoid is retinol which has to be bound to retinol-binding protein (RBP). RBP is the principle vitamin A carrier protein in the blood (Sun, 2012). Plasma levels of retinol are regulated homeostatically. Retinol functions via its active metabolites such as all-trans-retinoic acid that regulates gene transcription by activating the nuclear receptor retinoic acid receptor (RARs) (Chambon, 1996; Germain *et al.*, 2006; Schug *et al.*, 2007; Shaw *et al.*, 2003). Retinol uptake from blood is enabled by a protein called stimulated by retinoic acid 6 (STRA6) (Figure 2.34). STRA6 is a cell surface transporter which binds RBP and facilitates the movement of retinol from serum to cells from the vitamin A/RBP complex (holo-RBP) (Kawaguchi *et al.*, 2007; Kawaguchi *et al.*, 2008). It has also been proposed that STRA6 can function bidirectionally, being able to secrete the vitamin, as well as to take up retinol from the serum (Isken *et al.*, 2008).

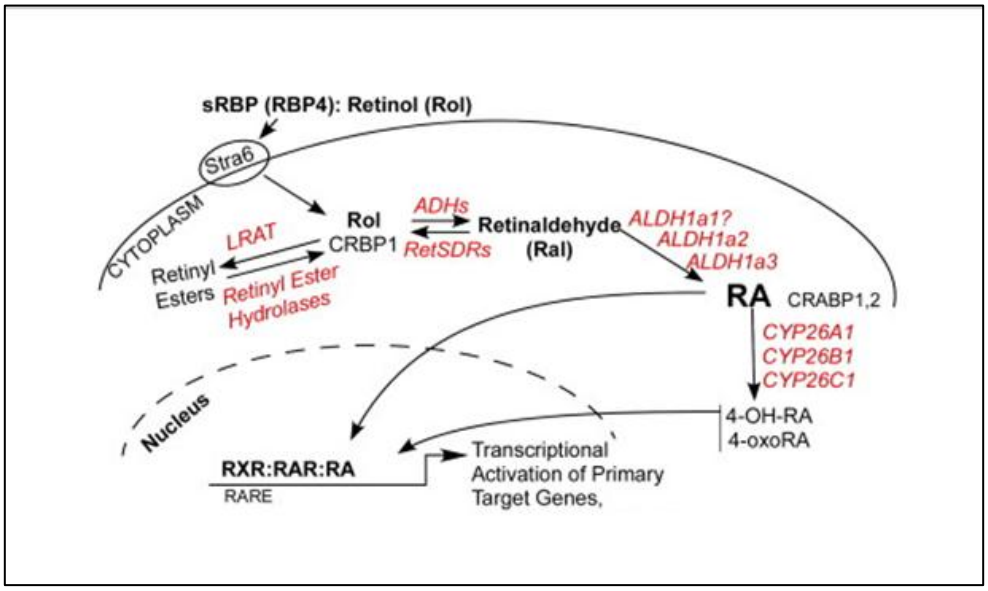


Figure 2.34: Summary of the retinoid signaling pathway and movement of retinol with its physiological carrier retinol binding protein (RBP) through STRA6. *STRA6 is a multi-transmembrane domain protein that binds to RBP and both Stra6 and lecithin: retinolacyltransferase (LRAT) facilitates retinol uptake. In the cytoplasm, alcohol dehydrogenase converts retinol into retinal. Retinal dehydrogenase (RALDH) then converts retinal to retinoic acid (RA). Retinoic acid is either located to the nucleus or stays in the cytoplasm, bound to cellular retinol binding protein (CRABP) 1 and 2. Retinoic acid receptor (RAR) in the nucleus interacts with retinoid X receptor (RXR) and binds to retinoic acid response element (RARE) upstream of the promoter region of a gene interest and leads transcription. In the cytoplasm, RA is broken down by Cyp26 enzymes (Guidas 2012).*

STRA6 was initially discovered in cancer cells as a retinoic acid promoted gene of unknown function. The likely explanation for the development of STRA6 in vertebrates is that these organisms use vitamin A for different biological activities, such as gene transcription regulation compared to other organisms. Gene expression of STRA6 is important for the regulation of the immune system, protecting against infection (Ross, 1992; Stephensen, 2001). Typically, STRA6 has nine transmembrane domains, with no similarity with any other membrane receptors, transporters or channels of known function (Kawaguchi *et al.*, 2007). There is no information on its substrate uptake mechanism because this does not depend on cellular energy (Kawaguchi *et al.*, 2007).

A vitamin A all-trans-retinol molecule is produced by two sequential oxidation steps; oxidation from retinol to retinaldehyde (RALD), and then RALD to RA. There are two retinol dehydrogenase enzyme classes, medium-chain alcohol dehydrogenases (ADH) and short-chain dehydrogenases (SDRs). Different short-chain dehydrogenases are distinctly related; where at least 20 separate enzymes are identified for the same protein family. Shared amino acid identity between enzyme pairs is typically around only 25% (Persson *et al.*, 1991; Theodosiou *et al.*, 2010).

Studies revealed that retinoids develop to inhibit thrombosis through increased fibrinolytic activity and decreased pro-thrombotic processes in vessel walls (Uruno *et al.*, 2003). It is known that vitamin A results in an increased inflammatory response (Wiedermann *et al.*, 1996). Therefore, considering the involvement of vitamin A in regulating the immune response; RA alters host sensitivity to infection by modulating both innate and adaptive immunity performance, as well as host inflammatory reactions (Mora *et al.*, 2008). Furthermore, RA increases cytotoxicity and T-cell proliferation, decreases B-cell proliferation and may also constrain B-cell apoptosis through RAR responses (Ballow *et al.*, 1996; Blomhoff *et al.*, 1992; Dennert, Lotan, 1978; Lomo *et al.*, 1998). Additionally, retinoic acids can alter antigen presenting by directly affecting dendritic cells (DC). For example, RA increases the expression of matrix metalloproteinases (Darmanin *et al.*, 2007).

STRA6, a vitamin A transporter, activates a signaling cascade assisted by the Janus kinases (JAKs), associated transcription factors and **s**ignal **t**ransducers and **a**ctivators of **t**ranscription (STATs). Pathway activation of JAK/STAT via extracellular peptides and their receptors causes the extracellular signal to alter the gene expression, regulating the various functions of proteins in the cells (Aaronson, Horvath, 2002; Darnell Jr *et al.*, 1994; Oshea *et al.*, 2002). The JAK-STAT system consists of three main components: (1) a receptor (2) Janus kinase (JAK) and a (3) signal transducer and activator of transcription (STAT). Most of the JAK-STAT pathways are found in white blood cells, and are therefore involved in regulation of the immune system. In addition both ALDHs and RDHs play roles in regulating the JAK/STAT cascade and protein kinase signaling cascade due to their oxidoreductase activity (Aaronson, Horvath, 2002).

Innate immune responses of arthropods are altered by the Toll, Immune efficiency and JAK/STAT systems (Fragkoudis *et al.*, 2009). Initially, the JAK/STAT system was discovered as a cytokine signaling pathway in mammals (Agaisse, Perrimon, 2004; Copf *et al.*, 2011). To date, it has been proven that it plays a critical role in defense mechanisms in human, *Drosophila* and mosquitoes immune responses (Ligoxygakis, 2012; Zou *et al.*, 2011). The JAK/STAT pathway has a role in controlling bacterial infection in *Drosophila* and *Plasmodium* invasion in mosquitoes (Bahia *et al.*, 2011). Liu *et al*. 2012, studied the *Ixodes scapularis* JAK/STAT pathway to identify regulation of tick antimicrobial peptides, which may aid in controlling the causative agent of human granulocytic anaplasmosis (Liu *et al.*, 2012). Their study proved the presence of the JAK/STAT pathway's involvement in *A. phagocytophilum* infection of ticks. Infection by *A. phagocytophilum* was drastically increased in salivary gland and the hemolymph of *I. scapularis* when the JAK/STAT pathway was inhibited using RNAi in addition to increased transmission of this pathogen to the mammalian host. Based on these results, it is assumed that JAK/STAT signaling plays a critical role in pathogen infection in ticks by affecting the expression of antimicrobial proteins (Liu *et al.*, 2012).

Several studies have been done to study the effect of Bm86 on pathogen transmission to the host as well as to the tick. The efficacy of a Bm86 vaccine has been tested in protection against ticks and pathogens transmitted by ticks (*Babesiabovis and B. bigemina*) (Pipano *et al.*, 2003). The results showed that

immunized cattle, when challenged with *B. bovis*-infected ticks, continued to become infected, however in the case of *B. bigemina* they became resistant to the infection. Bm86 is therefore suggested to be involved in pathogen transmission of certain species. The effect of Bm86 gene silencing on the fitness of *R. microplus* ticks fed on *B. bovis*-infected cattle was studied, leading to a decreased survival of engorged tick and egg weights (Bastos *et al.*, 2010). It is believed that Bm86 is involved with pathogen invasion within the tick, as well as from ticks to the host, since it was shown that Bm86 interacts with proteins that are part of a pathway that regulates the immune system.

2.5 CONCLUSION

The objective of this study was to investigate the possibility of identifying interacting proteins of Bm86 and BmATAQ from *R. microplus* by means of yeast two-hybrid screening. The objective extended to functional annotation of BmATAQ and Bm86 by analyzing their putative interacting partners to allow prediction of their respective biological roles, using bioinformatic tools. However, this is only possible if there is enough available data for the respective metabolic pathways they play a role in. This methodology has been shown to identify novel proteins with a number of biological functions in other organisms. In ticks, however, this technique has only been reported once for identification the function of Subolesin from *I. scapularis* (De la Fuente *et al.*, 2008). In this study, seven BmATAQ and four Bm86 putative binding candidates were identified using the yeast two-hybrid system.

The identified putative interacting candidates were further investigated to functionally annotate BmATAQ and Bm86. This was done using PSI-BLAST, followed by three dimensional modeling using threading. It was possible to identify the interacting proteins on amino acid level, but no hit could be obtained on nucleotide sequence level, since nucleotide sequences are less conserved. Structural sequence analysis was able to detect distinctly related proteins despite a sequence identity of less than 15%.

Two interacting partners were identified for both BmATAQ and Bm86. BmATAQ binding partners Clone 12 and clone 50 were found to be an ALDH and a RDH, respectively. Bm86 binding partners clone 53 and clone 93 were found to be a multi domain Kunitz-like protein and an ALDH, respectively. Confirmation of RDH as an interacting partner via Western blot analysis with an α-RDH antibody was done for both BmATAQ and Bm86. Hence, this study revealed that BmATAQ and Bm86 have common binding partners ALDH and RDH. In addition, a novel multi-domain Kunitz-like protein was found to interact with Bm86.

Both ALDHs and RDHs play critical roles in retinol and vitamin A metabolism with an oxidoreductase activity to regulate the immune response against pathogens. This correlates with the outcome of several studies that have been done to study the importance of Bm86 regarding pathogen transmission and tick survival. Commonly, these investigations proposed that Bm86 is involved in pathogen invasion and

altering immune responses to pathogens. Bm86 is a GPI-anchored gut protein and BmATAQ has a transmembrane and an intracellular domain. Both of these proteins are also abundantly expressed in the gut of adult female ticks. Based on these structural characteristics they both may be involved in cell signaling, specifically in immune responses.

Additionally a gene called Wfikkn1 encodes for a protein which also interacts with all-trans retinoic acids. This multi domain protein acts as a protease-inhibitor that contains multiple distinct protease inhibitor domains including Kunitz domain. It was hypothesized that this multiple domain protein has serine protease- and metalloprotease-inhibitor activity (Gibbs *et al.*, 2004; Kondas *et al.*, 2008; Nishida *et al.*, 2004). Since all three putative interacting partners (ALDH, RDH and Kunitz-like protein) of Bm86 and BmATAQ are involved in the same metabolic pathway one can hypothesize that both targeted proteins are involved in the same metabolic pathway.

This study can be concluded by purposing a pathway where Bm86, ATAQ and their predicted interacting partners functions together which was illustrated below (Figure 2.35).

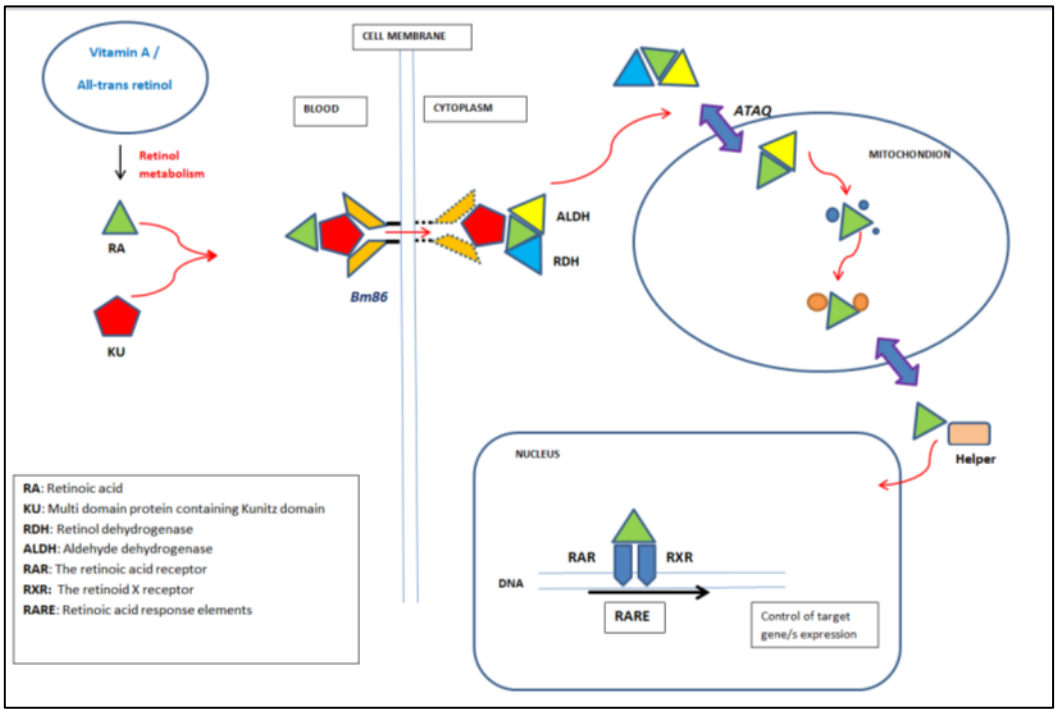


Figure 2.35: Schematic illustration of a proposed pathway of retinol metabolism which involves targeted Bm86, ATAQ and all putative interacting partners of them. *In the blood metabolized* ***all-trans retinol*** *might be taken into the cell in the form of retinoic acids (**RA**) with binding to**Bm86** with a protein has Kunitz domain (**KU**). RA could be bound with retinol dehydrogenases (**RDH**) and aldehyde dehydrogenases (**ALDH**) to become progressed and transported to the mitochondrion. **RA** might be transported into the mitochondrion with **ALDH** by interacting with **ATAQ** to progress further. **RA** enters into the nucleus and bind to retinoic acid receptor (**RAR**) and retinoic X receptor (**RXR**) on the DNA region called retinoic acid response element (**RARE**). This complex molecule then controls targeted gene/s expression, mostly to activate the transcription of the targeted genes.*

Considering the common pathway that both Bm86 and ATAQ are involved in would be the retinol metabolism (Figure 2.35). All-trans retinoic acids (RA) could be captured by a multi domain protein that includes Kunitz domains (KU). The KU protein brings RA close to Bm86 protein which is located on the cell membrane then the RA could be transported into the cell. In the cell RA bound by aldehyde dehydrogenases (ALDHs) and retinol dehydrogenases (RDHs) to process for

transporting to the mitochondrion. RA and ALDH could be transported into the mitochondrion with help from membrane bound ATAQ. RA processed through mitochondrion and transported to the nucleus. In the nucleus RA binds to retinoic acid receptor (RAR) and retinoic acid X receptor (RXR) located on the DNA in a region called retinoic acid response elements (RARE) (Parrado *et al.*, 2001). The RA, RAR and RXR complex is mostly involved in the control of transcription activation (Parrado *et al.*, 2001). Another proposed pathway for the same metabolism would be like illustrated in the figure below (Figure 2.36). All-trans retinoid in the blood meal is captured by KU and transported in to the cell through Bm86 and in the cell retinoid binds with retinol dehydrogenase (RDH). RDH catalyze retinoid into retinol aldehyde (RAL) which binds to ALDH and becomes RA. RA releases from ALDH and enters the mitochondrion via an interaction with ATAQ. After certain processes RA enters into the nucleus and impacts targeted gene/s expression.

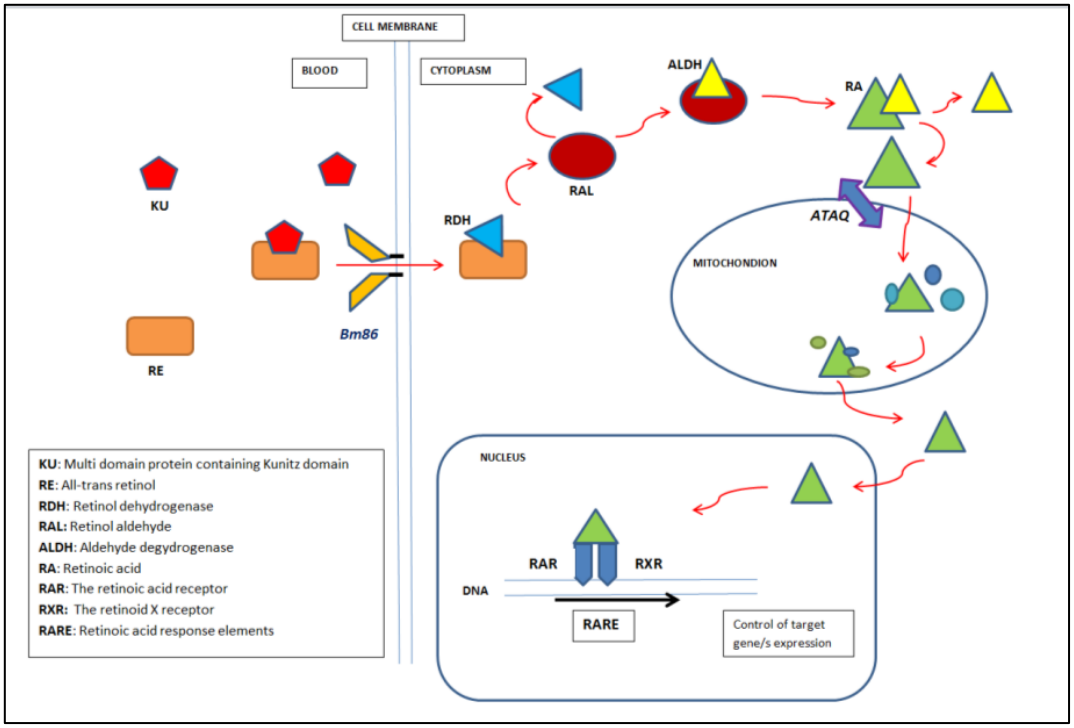


Figure 2.36: Schematic illustration of second proposed pathway of retinol metabolism in ticks which involves targeted Bm86, ATAQ and all putative interacting partners of them. *All-trans retinol (**RE**) might be captured by a protein has Kunitz domain (**KU**) and taken into the cell binding to **Bm86.** In the cell **RE** metabolized to retinoic acids (**RA**) with retinol dehydrogenases (**RDH**) and aldehyde dehydrogenases (**ALDH**) and transported to the mitochondrion via **ATAQ**. After RA progressed further in the mitochondrion it will enter into the nucleus to influence expression of targeted gene/s. In the nucleus **RA** binds to retinoic acid receptor (**RAR**) and retinoic X receptor (**RXR**) on the DNA region called retinoic acid response element (**RARE**) to activate transcriptions.*

2.6 REFERENCES

Aaronson DS, Horvath CM (2002) A road map for those who don't know JAK-STAT. *Science* **296**, 1653-1655.

Agaisse H, Perrimon N (2004) The roles of JAK/STAT signaling in *Drosophila* immune responses. *Immunological Reviews* **198**, 72-82.

Altschul SF, Boguski MS, Gish W, Wootton JC (1994) Issues in searching molecular sequence databases. *Nature Genetics* **6**, 119-129.

Altschul SF, Madden TL, Schaffer AA*, et al.* (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.

Andrade MA, Brown NP, Leroy C*, et al.* (1999) Automated genome sequence analysis and annotation. *Bioinformatics* **15**, 391-412.

Ascenzi P, Bocedi A, Bolognesi M*, et al.* (2003) The bovine basic pancreatic trypsin inhibitor (Kunitz inhibitor): A milestone protein. *Current Protein and Peptide Science* **4**, 231-251.

Bahia AC, Kubota MS, Tempone AJ*, et al.* (2011) The JAK-STAT pathway controls Plasmodium vivax load in early stages of Anopheles aquasalis infection. *PLoS Neglected Tropical Diseases* **5**(11): e1317.

Ballow M, Xiang S, Wang W, Brodsky L (1996) The effects of retinoic acid on immunoglobulin synthesis: Role of interleukin 6. *Journal of Clinical Immunology* **16**, 171-179.

Bastos RG, Ueti MW, Knowles DP, Scoles GA (2010) The *Rhipicephalus (Boophilus) microplus Bm86* gene plays a critical role in the fitness of ticks fed on cattle during acute *Babesia bovis* infection. *Parasites and Vectors* **3**:111.

Bergman CM, Quesneville H (2007) Discovering and detecting transposable elements in genome sequences. *Briefings in Bioinformatics* **8**, 382-392.

Björklund AK, Ekman D, Elofsson A (2006) Expansion of protein domain repeats. *PLoS Computational Biology* **2**, 0959-0970.

Blomhoff HK, Smeland EB, Erikstein B*, et al.* (1992) Vitamin A is a key regulator for cell growth, cytokine production, and differentiation in normal B cells. *Journal of Biological Chemistry* **267**, 23988-23992.

Bonifacino J. S. C. DAE (2001) Immunoprecipitation. *Current Protocols in Cell Biology* **114**, 2831-2841

Bonnet M, Kurz M, Mesa S*, et al.* (2013) The structure of *Bradyrhizobium japonicum* transcription factor FixK 2 unveils sites of DNA binding and oxidation. *Journal of Biological Chemistry* **288**, 14238-14246.

Bowie JU, Luthy R, Eisenberg D (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science* **253**, 164-170.

Carey MF, Peterson CL, Smale ST (2012) Confirming the functional importance of a protein-DNA interaction. *Cold Spring Harbor Protocols* **7**, 733-757.

Chambon P (1996) A decade of molecular biology of retinoic acid receptors. *FASEB Journal* **10**, 940-954.

Chothia C, Janin J (1975) Principles of protein-protein recognition. *Nature* **256**, 705-708.

Copf T, Goguel V, Lampin-Saint-Amaux A, Scaplehorn N, Preat T (2011) Cytokine signaling through the JAK/STAT pathway is required for long-term memory in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 8059-8064.

Cox J, Mann M (2011) Quantitative, high-resolution proteomics for data-driven systems biology, pp. 273-299.

Darmanin S, Chen J, Zhao S*, et al.* (2007) All-trans retinoic acid enhances murine dendritic cell migration to draining lymph nodes via the balance of matrix metalloproteinases and their inhibitors. *Journal of Immunology* **179**, 4616-4625.

Darnell Jr JE, Kerr IM, Stark GR (1994) Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* **264**, 1415-1421.

De la Fuente J, Maritz-Olivier C, Naranjo V*, et al.* (2008) Evidence of the role of tick subolesin in gene expression. *BMC Genomics* **9**:372.

De Las Rivas J, Fontanillo C (2010) Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS Computational Biology* **6**(6): e1000807.

Deng M, Zhang K, Mehta S, Chen T, Sun F (2002) Prediction of protein function using protein-protein interaction data. *Proceedings IEEE Computer Society Bioinformatics Conference* **1**, 197-206.

Dennert G, Lotan R (1978) Effects of retinoic acid on the immune system: stimulation of T killer cell induction. *European Journal of Immunology* **8**, 23-29.

Dunker AK, Lawson JD, Brown CJ*, et al.* (2001) Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* **19**, 26-59.

Elofsson A (2002) A study on protein sequence alignment quality. *Proteins: Structure, Function and Genetics* **46**, 330-339.

Fields S, Song O (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-247.

Fischer M, Zhang QC, Dey F*, et al.* (2011) MarkUs: A server to navigate sequence-structure-function space. *Nucleic Acids Research* **39**, W357-W361.

Fragkoudis R, Attarzadeh-Yazdi G, Nash AA, Fazakerley JK, Kohl A (2009) Advances in dissecting mosquito innate immune responses to arbovirus infection. *Journal of General Virology* **90**, 2061-2072.

Francischetti IMB, Mather TN, Ribeiro JMC (2004) Penthalaris, a novel recombinant five-Kunitz tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick vector of Lyme disease, *Ixodes scapularis*. *Thrombosis and Haemostasis* **91**, 886-898.

Francischetti IMB, Pham VM, Mans BJ*, et al.* (2005) The transcriptome of the salivary glands of the female western black-legged tick *Ixodes pacificus* (Acari: Ixodidae). *Insect Biochemistry and Molecular Biology* **35**, 1142-1161.

Gatzeva-Topalova PZ, May AP, Sousa MC (2004) Crystal structure of *Escherichia coli* ArnA (PmrI) decarboxylase domain. A key enzyme for lipid A modification with 4-amino-4-deoxy-L-arabinose and polymyxin resistance. *Biochemistry* **43**, 13370-13379.

Germain P, Chambon P, Eichele G*, et al.* (2006) International union of pharmacology. LX. Retinoic acid receptors. *Pharmacological Reviews* **58**, 712-725.

Gibbs RA, Weinstock GM, Metzker ML*, et al.* (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428**, 493-520.

Gillis J, Pavlidis P (2012) "Guilt by association" is the exception rather than the rule in gene networks. *PLoS Computational Biology* **8**(3): e1002444.

Harlow E, Lane D (2006) Immunoaffinity Purification: Coupling Antibodies to Activated Beads. *Cold Spring Harbor Protocols* **2006**, pdb.prot4304.

Hedbacker K, Carlson M (2008) SNF1/AMPK pathways in yeast. *Frontiers in Bioscience* **13**, 2408-2420.

Hu L, Huang T, Shi X*, et al.* (2011) Predicting functions of proteins in mouse based on weighted protein-protein interaction network and protein hybrid properties. *PLoS ONE* **6**(1): e14556.

Isken A, Golczak M, Oberhauser V*, et al.* (2008) RBP4 Disrupts Vitamin A Uptake Homeostasis in a STRA6-Deficient Animal Model for Matthew-Wood Syndrome. *Cell Metabolism* **7**, 258-268.

Janin J, Chothia C (1990) The structure of protein-protein recognition sites. *Journal of Biological Chemistry* **265**, 16027-16030.

Jones DT, Taylor WR, Thornton JM (1992) A new approach to protein fold recognition. *Nature* **358**, 86-89.

Jones S (2012) Computational and structural characterisation of protein associations. *Advances in Experimental Medicine and Biology* **747**, 42-54.

Kalanon M, McFadden GI (2008) The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: A bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. *Genetics* **179**, 95-112.

Kavanagh KL, Jornvall H, Persson B, Oppermann U (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families: The SDR superfamily: Functional and structural diversity within a family of metabolic and regulatory enzymes. *Cellular and Molecular Life Sciences* **65**, 3895-3906.

Kawaguchi R, Yu J, Hu J*, et al.* (2007) A membrane receptor for retinol binding protein mediates cellular uptake of vitamin A. *Science* **315**, 820-825.

Kawaguchi R, Yu J, Wiita P, Ter-Stepanian M, Sun H (2008) Mapping the membrane topology and extracellular ligand binding domains of the retinol binding protein receptor. *Biochemistry* **47**, 5387-5395.

Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols* **4**, 363-371.

Kim Y, Min B, Yi GS (2011) IDDI: The integrated domain-domain interaction analysis system. *Bioinformatics and Biomedicine, IEEE* 520-525.

Klenova E, Chernukhin I, Inoue T, Shamsuddin S, Norton J (2002) Immunoprecipitation techniques for the analysis of transcription factor complexes. *Methods* **26**, 254-259.

Kohn KW (1999) Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Molecular Biology of the Cell* **10**, 2703-2734.

Kondas K, Szlama G, Trexler M, Patthy L (2008) Both WFIKKN1 and WFIKKN2 have high affinity for growth and differentiation factors 8 and 11. *Journal of Biological Chemistry* **283**, 23677-23684.

Koonin EV (2003) Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nature Reviews. Microbiology* **1**, 127-136.

Levitt M, Gerstein M (1998) A unified statistical framework for sequence comparison and structure comparison. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 5913-5920.

Liden M, Eriksson U (2006) Understanding retinol metabolism: Structure and function of retinol dehydrogenases. *Journal of Biological Chemistry* **281**, 13001-13004.

Ligoxygakis P (2012) Genetics of Immune Recognition and Response in *Drosophila* host defense. In: *Advances in Genetics*, pp. 71-97.

Liu L, Dai J, Zhao YO*, et al.* (2012) Ixodes scapularis JAK-STAT pathway regulates tick antimicrobial peptides, thereby controlling the agent of human granulocytic anaplasmosis. *Journal of Infectious Diseases* **206**, 1233-1241.

Liu x, Ohta T, Kawabata T, Kawai F (2013) Catalytic mechanism of short ethoxy chain nonylphenol dehydrogenase belonging to a polyethylene glycol dehydrogenase

group in the GMC oxidoreductase family. *International Journal of Molecular Sciences*.

Liu ZJ, Sun YJ, Rose J*, et al.* (1997) The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD^+ and the Rossmann fold. *Nature Structural Biology* **4**, 317-326.

Lomo J, Smeland EB, Ulven S*, et al.* (1998) RAR-, not RXR, ligands inhibit cell activation and prevent apoptosis in B-Lymphocytes. *Journal of Cellular Physiology* **175**, 68-77.

Looso M, Borchardt T, Krüger M, Braun T (2010) Advanced identification of proteins in uncharacterized proteomes by pulsed in vivo stable isotope labeling-based mass spectrometry. *Molecular and Cellular Proteomics* **9**, 1157-1166.

Lu H, Skolnick J (2001) A distance-dependent atomic knowledge-based potential for improved protein structure selection. *Proteins: Structure, Function and Genetics* **44**, 223-232.

Macedo-Ribeiro S, Almeida C, Calisto BM*, et al.* (2008a) Isolation, cloning and structural characterisation of boophilin, a multifunctional Kunitz-type proteinase inhibitor from the cattle tick. *PLoS ONE* **3**:e1624.

Macedo-Ribeiro S, Almeida C, Calisto BM*, et al.* (2008b) Isolation, cloning and structural characterization of boophilin, a multifunctional Kunitz-type proteinase inhibitor from the cattle tick. *PLoS ONE* **3**, 1-17.

Maeda T, Yoshinaga I, Shiba T*, et al.* (2000) Cloning and sequencing of the gene encoding an aldehyde dehydrogenase that is induced by growing Alteromonas sp. strain KE10 in a low concentration of organic nutrients. *Applied and Environmental Microbiology* **66**, 1883-1889.

Marchitti SA, Brocker C, Stagos D, Vasiliou V (2008) Non-P450 aldehyde oxidizing enzymes: The aldehyde dehydrogenase superfamily. *Expert Opinion on Drug Metabolism and Toxicology* **4**, 697-720.

Maritz-Olivier C, van Zyl W, Stutzer C (2012) A systematic, functional genomics, and reverse vaccinology approach to the identification of vaccine candidates in the cattle tick, *Rhipicephalus microplus*. *Ticks and Tick-Borne Diseases* **3**, 179-187.

Matecic M, Smith Jr DL, Pan X*, et al.* (2010) A microarray-based genetic screen for yeast chronological aging factors. *PLoS Genetics* **6**(4): e1000921.

McCafferty J, Griffiths AD, Winter G, Chiswell DJ (1990) Phage antibodies: Filamentous phage displaying antibody variable domains. *Nature* **348**, 552-554.

McCollum M, Davis M (1913) The necessity of csertain lipins in the diet during growth. *The Journal of Biological Chemistry* **14**, 167-175.

Mészáros B, Simon I, Dosztányi Z (2009) Prediction of protein binding regions in disordered proteins. *PLoS Computational Biology* **5**.

Miernyk JA, Thelen JJ (2008) Biochemical approaches for discovering protein-protein interactions. *Plant Journal* **53**, 597-609.

Miller J, Stagljar I (2004) Using the yeast two-hybrid system to identify interacting proteins. *Methods in Molecular Biology (Clifton, N.J.)* **261**, 247-262.

Mora JR, Iwata M, Von Andrian UH (2008) Vitamin effects on the immune system: Vitamins A and D take centre stage. *Nature Reviews Immunology* **8**, 685-698.

Nesvizhskii AI (2012) Computational and informatics strategies for identification of specific protein interaction partners in affinity purification mass spectrometry experiments. *Proteomics* **12**, 1639-1655.

Nishida AT, Kobuke K, Kojima K*, et al.* (2004) OC29 is preferentially expressed in the presumptive sensory organ region of the otocyst. *Developmental Dynamics* **231**, 766-774.

Noy N (2006) Vitamin A. In: *Biochemical, Physiological & Molecular Aspects of Human Nutrition* (ed. Stipanuk MH). Saunder Elsevier, St. Loius.

Osborne TB, Mendel LB (1919) The vitamins in green foods. *The Journal of Biological Chemistry* **37**, 187-200.

Oshea JJ, Gadina M, Schreiber RD (2002) Cytokine signaling in 2002: New surprises in the Jak/Stat pathway. *Cell* **109**, S121-S131.

Paesen GC, Siebold C, Dallas ML*, et al.* (2009) An ion-channel modulator from the saliva of the brown ear tick has a highly modified Kunitz/BPTI structure. *Journal of Molecular Biology* **389**, 734-747.

Parrado A, Despouy G, Kraiba R*, et al.* (2001) Retinoic acid receptor α1 variants, RARα1ΔB and RARα1ΔBC, define a new class of nuclear receptor isoforms. *Nucleic Acids Research* **29**, 4901-4908.

Pazos F, Bang J-W (2006) Computational Prediction of Functionally Important Regions in Proteins. *Current Bioinformatics* **1**, 15-23.

Pearson WR (1996) Effective protein sequence comparison. *Methods in Enzymology**266***, 227-256.

Persson B, Krook M, Jornvall H (1991) Characteristics of short-chain alcohol dehydrogenases and related enzymes. *European Journal of Biochemistry* **200**, 537-543.

Phizicky EM, Fields S (1995) Protein-protein interactions: Methods for detection and analysis. *Microbiological Reviews* **59**, 94-123.

Pipano E, Alekceev E, Galker F*, et al.* (2003) Immunity against *Boophilus annulatus* induced by the Bm86 (Tick-GARD) vaccine. *Experimental and Applied Acarology* **29**, 141-149.

Rigaut G, Shevchenko A, Rutz B*, et al.* (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nature Biotechnology* **17**, 1030-1032.

Rigden DJ, Woodhead DD, Wong PWH, Galperin MY (2011) New structural and functional contexts of the Dx[DN]xDG linear motif: Insights into evolution of Calcium-binding proteins. *PLoS ONE* **6**(6): e21507.

Ross AC (1992) Vitamin A status: Relationship to immunity and the antibody response. *Proceedings of the Society for Experimental Biology and Medicine* **200**, 303-320.

Schaefer C, Schlessinger A, Rost B (2010) Protein secondary structure appears to be robust under in silico evolution while protein disorder appears not to be. *Bioinformatics* **26**, 625-631.

Schagger H, Von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry* **166**, 368-379.

Schug TT, Berry DC, Shaw NS, Travis SN, Noy N (2007) Opposing effects of retinoic acid on cell growth result from alternate activation of two different nuclear receptors. *Cell* **129**, 723-733.

Schuler GD (1998) Sequence alignment and database searching. *Methods Biochem Anal* **39**, 145-171.

Schuller HJ (2003) Transcriptional control of nonfermentative metabolism in the yeast *Saccharomyces cerevisiae*. *Current Genetics* **43**, 139-160.

Schwikowski B, Uetz P, Fields S (2000) A network of protein-protein interactions in yeast. *Nature Biotechnology* **18**, 1257-1261.

Serang O, Noble W (2012) A review of statistical methods for protein identification using tandem mass spectrometry. *Statistics and its Interface* **5**, 3-20.

Shaw N, Elholm M, Noy N (2003) Retinoic acid is a high affinity selective ligand for the peroxisome proliferator-activated receptor beta/delta. *Journal of Biological Chemistry* **278**, 41589-41592.

Sippl MJ, Flöckner H (1996) Threading thrills and threats. *Structure* **4**, 15-19.

Skolnick J, Fetrow JS (2000) From genes to protein structure and function: Novel applications of computational approaches in the genomic era. *Trends in Biotechnology* **18**, 34-39.

Sleator RD (2012) Prediction of protein functions. In: *Methods in Molecular Biology*, pp. 15-24.

Smith GP (1985) Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-1317.

Stephensen CB (2001) Vitamin A, infection, and immune function. In: *Annual Review of Nutrition*, pp. 167-192.

Sun H (2012) Membrane receptors and transporters involved in the function and transport of vitamin A and its derivatives. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1821**, 99-112.

Theodosiou M, Laudet V, Schubert M (2010) From carrot to clinic: An overview of the retinoic acid signaling pathway. *Cellular and Molecular Life Sciences* **67**, 1423-1445.

Torda AE (1997) Perspectives in protein-fold recognition. *Current Opinion in Structural Biology* **7**, 200-205.

Uruno A, Sugawara A, Kudo M*, et al.* (2003) Transcription suppression of thromboxane receptor gene expression by retinoids in vascular smooth muscle cells. *Hypertension Research* **26**, 815-821.

Vasiliou V, Pappa A, Petersen DR (2000) Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chemico-Biological Interactions* **129**, 1-19.

Vasiliou V, Sandoval M, Backos DS*, et al.* (2013) ALDH16A1 is a novel non-catalytic enzyme that may be involved in the etiology of gout via protein-protein interactions with HPRT1. *Chemico-Biological Interactions* **202**, 22-31.

Vinayagam A, Stelzl U, Foulle R*, et al.* (2011) A directed protein interaction network for investigating intracellular signal transduction. *Science Signaling* **4**.

Wiedermann U, Chen XJ, Enerbäck L*, et al.* (1996) Vitamin A deficiency increases inflammatory responses. *Scandinavian Journal of Immunology* **44**, 578-584.

Wierenga RK, Terpstra P, Hol WGJ (1986) Prediction of the occurrence of the ADP-binding beta alpha beta-fold in proteins, using an amino acid sequence fingerprint. *Journal of Molecular Biology* **187**, 101-107.

Wilson N (2004) Interacting maps. *Nature Reviews Molecular Cell Biology* **5**, 80-80.

Xia H, Mathew B, John T, Hegab H, Feng J (2013) Microfluidic immunoprecipitation for post-translational modified protein purification **39 IFMBE**, 297-300.

Yang H, Zhang D, Wang J, Wood AJ, Zhang Y (2012) Molecular cloning of a stress-responsive aldehyde dehydrogenase gene ScALDH21 from the desiccation-tolerant moss *Syntrichia caninervis* and its responses to different stresses. *Molecular Biology Reports* **39**, 2645-2652.

Young KH (1998) Yeast two-hybrid: So many interactions, (in) so little time. *Biology of Reproduction* **58**, 302-311.

Young RA, Davis RW (1983) Yeast RNA polymerase II genes: Isolation with antibody probes. *Science* **222**, 778-782.

Zou Z, Souza-Neto J, Xi Z*, et al.* (2011) Transcriptome analysis of *Aedes aegypti* transgenic mosquitoes with altered immunity. *PLoS Pathogens* **7**(11): e1002394.

Chapter 3: Concluding discussion

Rhipicephalus microplus belongs to the suborder Ixodidae and is a vector for a wide range of pathogens such as *Anaplasma marginale* (anaplasmosis), *Babesia bigemina* and *B. bovis* (babesiosis). The resultant diseases are important as they lead to major economic losses (Jongejan, Uilenberg, 2004). Furthermore, the medical importance of *R. microplus* is further highlighted by the recent increase in the prevalence of zoonotic diseases such as babesiosis (Perez de Leon *et al.*, 2010).

Harsh chemical control of ticks leads to contamination of milk and meat products, as well as the environment. Moreover, there has been a global increase in the incidence of acaricide-resistant ticks (Parizi *et al.*, 2009). Immunological control presents a more economical and environmentally-friendly alternative to chemical control of ticks (Almazán *et al.*, 2010). However, only one vaccine against tick infestation has been commercialized and it is derived from a concealed antigen, Bm86, isolated from *R. microplus*. The efficacy of this vaccine varies against different tick species and geographical strains, where the average effect of the vaccine is 50% ranging from 0 to 100% (De Vos *et al.*, 2001). Considering that Bm86 is the only available anti-tick vaccine, it is hypothesized that improving the vaccine might be a viable option.

An improvement of current vaccines could be achieved by identifying homologs of the current antigen. In this regard, a novel protein ATAQ was identified during a study that aimed to identify Bm86 homologs. It was shown that these two proteins do not share identical antigenic and immunogenic regions and that they also have different expression patterns (Nijhof *et al.*, 2010; Patarroyo *et al.*, 2002). This indicates that ATAQ is not a Bm86 homolog, but rather an entirely different protein. ATAQ is abundantly expressed in the gut and Malpighian tubules, whereas Bm86 is predominantly expressed in the gut. However, vaccine trials did not confer protection against tick infestation using a recombinant ATAQ vaccine (Dr. A. Nijhof, personal communication).

In this study it was hypothesized that Bm86 and ATAQ interacting partners can be discovered by using the yeast two-hybrid system. Additionally, functional annotation

of the Bm86 and ATAQ interacting proteins might provide insight into their biological functions which will lead to the development of a more efficient anti-tick vaccine.

During this study it was shown that the yeast two-hybrid system is an applicable technique to discover binding partners of a tick protein of interest. Only one other example exists in literature where this technique was used successfully in ticks (De la Fuente *et al.*, 2008). In this study, the yeast-two hybrid system followed by sequence and structural analyses, led to the identification of two interacting partners for Bm86: an aldehyde dehydrogenase (ALDH) and a multi domain Kunitz-like protein. Similarly, two interacting partners were found for ATAQ: an ALDH and a retinol dehydrogenase (RDH). The existence of the multi-domain Kunitz-like protein and RHD as binding partners was confirmed by performing searches against a protein database with PSI-BLAST, as well as against an in-house assembled tick database (Maritz-Olivier *et al.*, 2012). Furthermore, it was also found that Bm86 and ATAQ have a common interacting partner, ALDH, which was proven on amino acid level, as well as structural level. Western blot analysis showed that RDH is also a common binding partner between ATAQ and Bm86. Hence, two common interacting partners and a novel, Kunitz-like protein, were identified in this study, from which it can be hypothesized that Bm86 and ATAQ are involved in the same metabolic pathway/s.

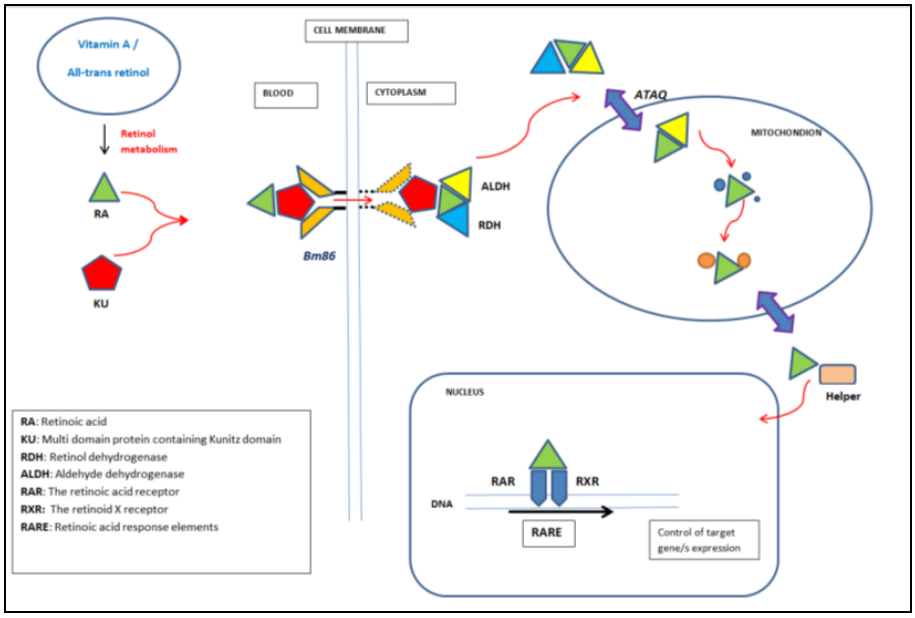
3.1. The retinoid pathway in ticks

Aldehyde and retinol dehydrogenases have oxidoreductase activity and are involved in the metabolism of retinoids (vitamin A). The end products of this pathway, retinoic acids (RAs), are part of a wide range of biological functions in almost all tissues (Sandell *et al.*, 2012). One of the RA biological roles is as part of the immune response, where RA can alter both the innate and adaptive immune responses (McGhee, Fujihashi, 2012). This might be an explanation for how changes in expression of Bm86 modulate pathogen infection in both the host organism and the tick itself.

The predictions made from this study also suggest that it is most likely that Bm86 plays a role in immune responses against pathogen invasion as proposed by literature (Brake *et al.*, 2010). This study can be extended further to look for more interacting partners of other anti-tick vaccine candidates to develop an anti-tick vaccine with improved efficacy.

In future studies, other components of the retinol metabolic pathway can be identified by looking at similar transcripts using available *R. microplus* sequence databases and a real-time PCR analysis of these transcripts in midgut tissues. Another approach can involve the application of DNA microarrays where Bm86 would be knocked-down *in vivo* and the resultant gene regulation analyzed for known components in the purposed pathway. To further confirm the predicted structures of putative interacting partners, expression of these proteins could be performed for crystallization and solution structure determination, as well as possible activity studies to confirm function. Finally, cellular localization studies can also be performed to confirm where these proteins occur *in situ*. In this regard, fluorescent microscopy using specific antibodies raised against recombinant proteins could be applied.

Taken together, the results of this study allows us to propose a possible metabolic pathway, where both Bm86 and ATAQ and their interacting partners function together (Figure 3.1 A and B). Since we predicted that the interacting partners of both Bm86 and ATAQ are ALDH, RDH and Kunitz-like, it is assumed that the metabolic pathway they involved is retinol metabolic pathway (Chapter 2). A multi domain Kunitz-like protein aids capture of All-trans retinoic acids (RA) and delivery to Bm86 protein which is located on the cell membrane to be transported into the cell. In the cell RA bound by aldehyde dehydrogenases (ALDHs) and retinol dehydrogenases (RDHs) to process for transporting to the mitochondrion. RA and ALDH could be transported into the mitochondrion with help from membrane bound ATAQ. RA processed through mitochondrion and transported to the nucleus. In the nucleus RA binds to retinoic acid receptor (RAR) and retinoic acid X receptor (RXR) located on the DNA in a region called retinoic acid response elements (RARE) (Parrado *et al*., 2001). The RA, RAR and RXR complex is mostly involved in the control of transcription activation (Parrado *et al*., 2001).



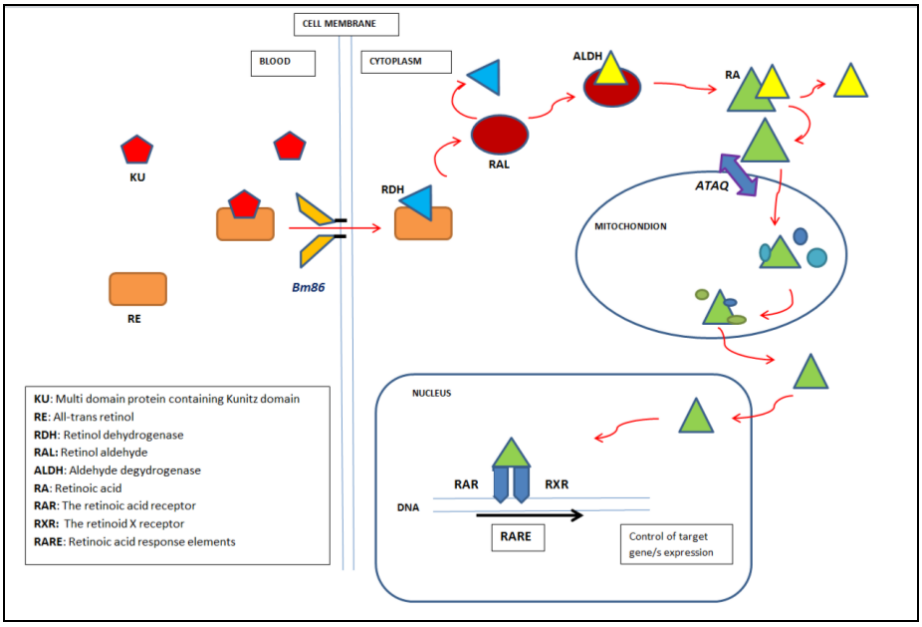


Figure 3.1: Proposed pathways of *R. microplus* retinol metabolism which involves targeted Bm86, ATAQ and all putative interacting partners of them. ***Pathway A:*** *In the blood metabolized* ***all-trans retinol*** *might be taken into the cell in the form of retinoic acids (**RA**) with binding to **Bm86** with a protein has Kunitz domain (**KU**). RA could be bound with retinol dehydrogenases (**RDH**) and aldehyde dehydrogenases (**ALDH**) to become progressed and transported to the mitochondrion. **RA** might be transported into the mitochondrion with **ALDH** by interacting with **ATAQ** to progress further. **RA** enters into the nucleus and bind to retinoic acid receptor (**RAR**) and retinoic X receptor (**RXR**) on the DNA region called retinoic acid response element (**RARE**). This complex molecule then controls targeted gene/s expression, mostly to activate the transcription of the targeted genes.* ***Pathway B:*** *All-trans retinol (**RE**) might be captured by a protein has Kunitz domain (**KU**) and taken into the cell binding to **Bm86.** In the cell **RE** metabolized to retinoic acids (**RA**) with retinol dehydrogenases (**RDH**)*

*and aldehyde dehydrogenases (**ALDH**) and transported to the mitochondrion via **ATAQ**. After RA progressed further in the mitochondrion it will enter into the nucleus to influence expression of targeted gene/s. In the nucleus **RA** binds to retinoic acid receptor (**RAR**) and retinoic X receptor (**RXR**) on the DNA region called retinoic acid response element (**RARE**) to activate transcriptions.*

In conclusion, interacting partners of Bm86 and ATAQ were successfully identified during this study. The next step would be to investigate the efficacy of a vaccine containing a mixture of both Bm86 and its interacting partners. It is hypothesized that such a combinatorial vaccine would have a higher efficacy against tick infestation. Moreover, having multiple targets in a metabolic pathway where the products involved in transcription activation might be more destructive in terms of tick survival.

REFERENCES

Almazán C, Lagunes R, Villar M*, et al.* (2010) Identification and characterization of *Rhipicephalus (Boophilus) microplus* candidate protective antigens for the control of cattle tick infestations. *Parasitology Research* **106**, 471-479.

Brake DK, Wikel SK, Tidwell JP, Pérez De León AA (2010) *Rhipicephalus microplus* salivary gland molecules induce differential CD86 expression in murine macrophages. *Parasites and Vectors* **3**:103.

De la Fuente J, Maritz-Olivier C, Naranjo V*, et al.* (2008) Evidence of the role of tick subolesin in gene expression. *BMC Genomics* **9**:372.

De Vos S, Zeinstra L, Taoufik A, Willadsen P, Jongejan F (2001) Evidence for the utility of the Bm86 antigen from *Boophilus microplus* in vaccination against other tick species. *Experimental and Applied Acarology* **25**, 245-261.

Jongejan F, Uilenberg G (2004) The global importance of ticks. *Parasitology* **129**, S3-S14.

Maritz-Olivier C, van Zyl W, Stutzer C (2012) A systematic, functional genomics, and reverse vaccinology approach to the identification of vaccine candidates in the cattle tick, *Rhipicephalus microplus*. *Ticks and Tick-Borne Diseases* **3**, 179-187.

McGhee JR, Fujihashi K (2012) Inside the mucosal immune system. *PLoS Biology* **10**(9): e1001397.

Nijhof AM, Balk JA, Postigo M*, et al.* (2010) Bm86 homologues and novel ATAQ proteins with multiple epidermal growth factor (EGF)-like domains from hard and soft ticks. *International Journal for Parasitology* **40**, 1587-1597.

Parizi LF, Pohl PC, Masuda A, Vaz Junior IS (2009) New approaches toward anti-*Rhipicephalus (Boophilus) microplus* tick vaccine. *Novas estrategias para o desenvolvimento de umavacina contra o carrapato Rhipicephalus (Boophilus) microplus. Revista Brasileira de Parasitologia Veterinaria* **18**, 1-7.

Patarroyo JH, Portela RW, De Castro RO*, et al.* (2002) Immunization of cattle with synthetic peptides derived from the *Boophilus microplus* gut protein (Bm86). *Veterinary Immunology and Immunopathology* **88**, 163-172.

Perez de Leon A, Strickman D, Knowles D*, et al.* (2010) One health approach to identify research needs in bovine and human babesioses: workshop report. *Parasites & Vectors* **3**, 36.

Sandell LL, Lynn ML, Inman KE, McDowell W, Trainor PA (2012) RDH10 oxidation of vitamin A is a critical control step in synthesis of retinoic acid during mouse embryogenesis. *PLoS ONE* **7**.

Appendix 1: Nucleotide sequence alignment of Cln13 and Cln12 (putative aldehyde dehydrogenase- predicted Bm86 interacting partner)

```
cln13           ACCCATCAACCACTTTGTACAAGAAAGCTGGGTCGAATTCGCCCTTAAGCAGTGGTATCA
cln12           ACCCATCAACCACTTTGTACAAGAAAGCTGGGTCGAATTCGCCCTTAAGCAGTGGTATCA
                ************************************************************

cln13           ACGCAGAGTGGCCATGGAGGCCGGCAATTCGGTGGTCGTGCTCGCCAGTTCGGAGTGGCC
cln12           ACGCAGAGTGGCCATGGAGGCCGGCAATTCGGTGGTCGTGCTCGCCAGTTCGGAGTGGCC
                ************************************************************

cln13           GCTTCCGGCACTGCACTTCTGTCAGATCCTGCAGGCATCGGACATACCGCCGGGCGTGGT
cln12           GCTTCCGGCACTGCACTTCTGTCAGATCCTGCAGGCATCGGACATACCGCCGGGCGTGGT
                ************************************************************

cln13           CAATGTCCTGTCGGCCAAGGAGCAGCCGACGCTTCTACGCACGCTTGCAGAGCATCACGA
cln12           CAATGTCCTGTCGGCCAAGGAGCAGCCGACGCTTCTACGCACGCTTGCAGAGCATCACGA
                ************************************************************

cln13           CGTTGCCGCCATCTGGTACTATGGGGAGCAGGACAGCACGGACGCATTCCTCAGATGTGC
cln12           CGTTGCCGCCATCTGGTACTATGGGGAGCAGGACAGCACGGACGCATTCCTCAGATGTGC
                ************************************************************

cln13           CATCTCGCACAACGACAAACACTGCTGGGTGTTCCCTCTGAGAGACGACCAAGACCGCTA
cln12           CATCTCGCACAACGACAAACACTGCTGGGTGTTCCCTCTGAGAGACGACCAACACCGCTA
                ************************************************** *******

cln13           CTGGATGGGCTGGCAGATTGCCCAGCACGCGACTGTGCCCAAGAGCATCTGGCTACCGTT
cln12           CTGGATGGGCTGGCAGATTGCCCAGCACGCGACTGTGCCCAAGAGCATCTGGCTACCGGT
                ************************************************************ *

cln13           TGGCGATACCTTCGCTAACTGACGGCAAGGGTCCTTGGAAACGCC-CCGTCGCCCACACT
cln12           TGGCGATACCTTCGCTAACTGACGGCAAGGGTCCTTGGAAACGCAGCCGTCGGCCACACT
                ******************************************  ****** *******

cln13           GCCATGGGAGACCTTGATTATGAAAGCTTCACTTTAATTTACTTAAAGGGGCCCTGCAGC
cln12           GCCATGGGAGACCTTGATTATGAAAGCTTCACTTTAATTTACTTAAAGGGGCCCTGCAGC
                ************************************************************

cln13           TCTTATTGAGCATTGTC-CAAAACGCTGCGAATCACATTGGACGAGGCTCCCGAGAACAC
cln12           TTTTATTGAGCATTGTCAAAAAACGCTGCTAATCATATTGGTTGAGGCTCCCGAGAACAC
                *.***************  ********** *****.***** .*****************

cln13           GCGAGCCCAACAT-----------------------------------------------
cln12           GCGAGCCAAATATTACAGGACAGCACGTGGTCTACCATTCACAATAAATTGTCGAAGTCA
                ******* **.**

cln13            ------------------------------------------------------------
cln12           GCTAAATATTGCTTTCTCTTCTCTCGACAAGTGACAGAGGAAGCTCAAAAATCACTCGTA

cln13            ------------------------------------------------------------
-
cln12           GAAGGCTATCTATCAGCCATTGGCTGATTTGAACATGGCGCGCTCGGTCGTTACAGAGAT

cln13            ------------------------------------------------------------
cln12           CGCCGTGTGAGGCTGCGACTTGTCGATGCGTGCGCGCGCGATCACTCGAAAAGTCACGAG
```

```
cln13            ------------------------------------------------------------
cln12          AAGGAAAGAATAGAGACATGCAATTTCTTTTTTGTTAAAAACCCCCCTTCTAGGGAAAAT

cln13            ------------------------------------------------------------
cln12          AAAACCGGATTTTTATACACCCTTGTTTGGTAGGAGGGATACTATGGCTCACTGCGGACA

cln13            ------------------------------------------------------------
cln12          TTGAGTATACTGGGACTTGTAGGGGAACTCCTTGGCCCGCCTCCCCTTATATTGTACAAA

cln13            ------------------
cln12           ATATAAGAACAAAAAGA
```

APPENDIX 2: Amino acid alignment of clone 12 (predict Bm86 binding partner-aldehyde dehydrogenase) and hits from tick database (BMGI). *The glutamic acid active site is indicated with blue and the active residue is indicated with red. The cysteine active site is indicated in yellow and the active residue is indicated in green.*

```
TC16791         MDASDRGILLNRLADLIERDRCLIASLETLDNGKPFVDAYNIDLPLVIKCLRYYAGYADK
TC24916         ------------------------------------------------------------
CLN12           ------------------------------------------------------------

TC16791         NHGKTIPLDGSYFAYTRHEPVGVCGQIIPWNFPALMQAWKLGPALAMGNTVVMKPAEQTP
TC24916         -----------------------------------MQAWKLGPALAMGNTVVMKPAEQTP
CLN12           ------------------------------------------MEAGNSVVVLASSEWP
                                                           :  **:**: .:.: *

TC16791         LSALHVASLVAEAGFPPGVVNVVPGM-GPTAGAAVAGHKDVDKIAFTGSTEVGQQVMETA
TC24916         LSALHVASLVAEAGFPPGVVNVVPGM-GPTAGAAIAAHRDVDKIAFTGSTEVGQQVMETA
CLN12           LPALHFCQILQASDIPPGVVNVLSAKEQPTLLRTLAEHHDVAAIWYYGEQDSTDAFLRCA
                *.***...::  :.:*******:..   **   ::* *:**  * : *. :  : .:. *

TC16791         AKSNLKKVTLELGGKSPNIVFKDANLDEAIETSHFGLFFNQGQCCCAGSRIFVEGAIYDD
TC24916         AKSNLKKVTLELGGKSPNIVFKDANLDEAIETSHFGLFFNQGQCCCAGSRIFVEGAIYED
CLN12           ISHNDK-------------------------------------HC----------------
                 . * *                                     :*

TC16791         FVARSVALAKQRVLGDPFDANTTQGPQVDQEQLGKILTLIDSGKSEGARLLCGGARHGSR
TC24916         FVARSVTLAKQRVLGDPFDANTTQGPQVDQEQLGKILTLIDSGKSEGARLLCGGARHGSR
CLN12           ------------------------------------------------------------

TC16791         GYFVEPTVFSDVKDGMRIAREEIFGPVMQILRFENIGELIERANRTEYGLAASLFTQDLE
TC24916         GYFVEPTVFSDVKDGMRIAREEIFGPVMQILRFENIGELIERANQTKYGLAASLFTQDLE
CLN12           --WVFPL-----------------------------------RDDQHRY------------
                  :* *                                     * :: .*

TC16791         KALHFSSGIKAGTVWVNCYDVLTAQVPFGGYKMSGIGRELGEYGLEAYTQVKSVIMKIGQ
TC24916         KALHFSSGIKAGTVWVNCYDVLTAQVPFGGYKMSGIGRELGEYGLEAYTQVKSVIMKIGQ
CLN12           --------------WMGWQIAQHATVP-----------------------KSIWLPVGD
                             *:.   .  * **                        **: : :*:

TC16791         K-NS
TC24916         K-NS
CLN12           TFAN
                .  .
```