A reverse genetics approach to evaluate Metzincins as anti-*Rhipicephalus microplus* tick vaccine candidates

by

Annette-Christi Barnard

Submitted in partial fulfillment of the requirements for the degree

Philosophiae Doctor

in the Faculty of Natural and Agricultural Sciences
Department Biochemistry
University of Pretoria
Pretoria

March 2013

ACKNOWLEDGEMENTS

I am exceptionally grateful towards the following:

- Dr. C. Maritz-Olivier, at the Deparment of Genetics, University of Pretoria, as supervisor of this project, for her guidance, endless encouragement and scientific insight. Thank you for instilling a passion for molecular science in me.

- Prof. A.W.H. Neitz and Dr. A.R.M. Gaspar my co-supervisors, at the Department of Biochemistry, University of Pretoria, for their valuable advice and helpful suggestions.

- Prof. F. Jongejan, at the Utrecht Centre for Tick borne diseases (UCTD), University of Utrecht, The Netherlands, for his generosity of opening his lab, enabling me to complete an important part of my studies.

- Dr. A.M. Nijhof, for his tireless advice, support and insightful ideas during my research visit to The Netherlands and thereafter.Thank you for enabling me to experience Utrecht at its best, on a bycicle.

- My fellow students, whom I have worked with throughout my studies. Thank you for all your support during exciting and not so exciting lab findings. Thank you for numerous scientific and philosophical discussions. My special thanks goes to Elizabeth Louw, for sharing her admirable knowledge, your input was invaluable.

- The University of Pretoria and the National Research Foundation for financial assistance and for international travel grants.

- My parents, Daniel B. and Riana Badenhorst, for their endless love, their infinite patience, encouragement, prayer and support. You have taught me to be strong and to never give up hope.

- Family and friends, for all your love, motivation and prayer. A special thank you for everyone who listened when I needed someone to share joy as well as frustration.

- My husband, Paul S. Barnard, for your never-ending patience, prayer, love and belief in me. You are my inspiration and make the journey worth living.

- God Almighty, I stand in awe of your supremacy. I thank you for awaking a curiosity in my soul to seek Your wonders through research.

"The one who gets wisdom loves life; the one who cherishes understanding will soon prosper."
Proverbs 19.8

DECLARATION

I, Annette-Christi Barnard, declare that the thesis entitled “A reverse genetics approach to evaluate Metzincins as anti-*Rhipichephalus microplus* tick vaccine candidates” which I hereby submit for the degree of *Philosophiae Doctor* Biochemistry 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. Where secondary material is used, this has been carefully acknowledged and referenced in accordance with university requirements. I am aware of university policy and implications regarding plagiarism.

Signature: .....

Date: .....

SUMMARY

Tick proteins functioning in vital physiological processes such as blood meal uptake, digestion and reproduction are potential targets for anti-tick vaccines, since vaccination could disrupt these essential functions and ultimately affect tick survival. In this study we identified metzincin metalloproteases from *R. microplus,* the world's most economically important external ectoparasite of cattle, as potential vaccine candidates since they are implicated to be essential to blood-cavity formation, bloodmeal digestion and reproduction in ixodid ticks. A vaccine derived from a single member of such a large family pose the obstacle of redundancy within the family, that may allow the function of the targeted family member to be taken up by other family members. Therefore the aim of this study was not only to focus on the physiological importance of each metzincin transcript, but also to investigate the differential gene expression network between the different metzincin family members.

Eight transcripts encoding proteins containing the characteristic metzincin zinc-binding motif HEXXHXXG/NXXH/D and a unique methionine-turn were identified from native and in-house assembled *R. microplus* Expressed Sequence Tag databases. These were representative of five reprolysin-like and three astacin-like metzincin metalloproteases. Reverse transcription-PCR indicated that the reprolysins were most abundantly expressed in the salivary glands, whereas the astacins were most abundant in the midgut and ovaries. *In vivo* gene silencing utilizing RNA interference, was performed to assess a possible phenotype in silenced adult female *R. microplus* ticks during blood feeding and reproduction. RNAi against two reprolysins and one astacin significantly affected average egg weight as well as the oviposition rate. Moreover, integrated real time-PCR studies revealed an extensive cross organ network between the *R. microplus* metzincin transcripts, supporting the use of a combinatorial metzincin-based anti- *R. microplus* vaccine targeting multiple members of the large metzincin clan simultaneous. To conclusively evaluate the vaccination potential of the three identified metzincin candidates, the immunogenicity and protective properties of the recombinant proteins needs to be determined. Due to metalloproteases destructive activity and characteristics such as cysteine rich domains, only selected domains of the three candidates were expressed, using a cost effective *Eshericia coli* based expression system. Finally, the ability of each successfully expressed domain to elicit an immune response and serve as a protective antigen against *R. microplus* will be screened during vaccination trials in cattle.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1
Literature Review

Chapter 2
Identification, cloning and expression profiling of astacin and reprolysin homologues from the hard tick, *R. microplus*

Chapter 3

***In vivo* gene silencing of *Rhipicephalus microplus* metzincins**

Chapter 4

Recombinant protein expression for vaccination

LIST OF TABLES

Chapter 1
Literature Review

Chapter 2
Identification, cloning and expression profiling of astacin and reprolysin homologues from the hard tick, *R. microplus*

Chapter 3
In vivo gene silencing of *Rhipicephalus microplus* metzincins

Chapter 4
Recombinant protein expression for vaccination

ABBREVIATIONS

Å	Angstrom
A	adenosine/ alanine
ADAM	a disintegrin and metalloprotease
ADAMTS	ADAM containing trombospondin-like motifs
ADP	adenosine diphosphate
Ago	argonaute
AsC	AsContig
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
BmiGI	a database of cDNAs expressed in *Boophilus microplus*
bp	basepair
°C	degrees Celcius
C/Cys	cysteine
cDNA	complementary DNA
C-terminal	carboxy terminal
cfu	colony forming units
D/Asp	aspartic acid
dddH2O	double distilled deionised water
DTT	dithiothreitol
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dsDNA	double stranded DNA
dsRNA	double stranded RNA
E/Glu	glutamic acid
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
ELI	expression library immunisation
EGF	epidermal growth factor-like domain
EST	expressed sequence tag
ExPASy	expert protein analysis system
F/Phe	phenylalanine
G/Gly	glycine
Gbp	giga basepairs
GPI	glycosylphosphatidylinisotol
GSP	gene specific primer
H/His	histidine

K/Lys	lysine
kDa	kilo Dalton
*lac*Z	β-Galactosidase gene
LB	Luria-Berthani
LC-MS-MS	liquid chromatography-tandem mass spectrometry
M/Met	methionine
MCS	multiple cloning site
MG	midgut
MMP	matrix metalloprotease
MP	metalloprotease
mRNA	messenger RNA
MSP1α	major surface protein 1α
NCBI	National Centre for Biotechnology Information
N-terminal	amino terminal
O	ovaries
P/Pro	proline
PAZ	Piwi/Argonaute/Zille
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
qPCR	real-time PCR
R/Arg	arginine
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNAi	RNA interference
rpm	rotations per minute
RT	reverse transcriptase
S/Ser	serine
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SG	salivary glands
shRNA	short hairpin RNA
SID-1	systemic RNA interference deficient-1
siRNA	small interfering RNA
ssRNA	single stranded RNA
SVMP	snake venom metalloprotease
TAE	tris-acetate EDTA
TFPI	tissue factor pathway inhibitor
TIMP	tissue inhibitor of a metalloprotease
tRNA	transfer RNA

V/Val	valine
Y/Tyr	tyrosine

Chapter 1

Literature Review

1.1 Introduction

Ticks are obligate haematophagous ectoparasites, which are distributed worldwide from tropical regions to the Arctic. The impact of these arthropods as disease vectors on human well-being is second only to that of mosquitoes and their effect on animal hosts – including domestic animals, livestock and wildlife – is immeasurably greater (Jongejan and Uilenberg, 2004). Ticks transmit a greater variety of infectious agents than any other arthropod group. The diversity of the pathogenic organisms includes: protozoa (eg. *Babesia* and *Theileria spp.*), rickettsiae (eg. *Anaplasma* and *Ehrlichia spp.*), bacteria (eg. *Borrelia* and *Coxiella spp.*), fungi (eg. *Scopulariopsis spp.*) and viruses (e.g. *Tick-Borne Encephalitis Virus*) (Jongejan and Uilenberg, 2004). Besides the pathogens they indirectly transmit, ticks also induce direct conditions such as toxicosis, anaemia and paralysis (Sonenshine, 1991). Moreover, due to their ferocious pool feeding habit ticks cause direct harm to their hosts' hides and promote secondary infections. Failure to control tick burdens results in great economical losses worldwide, since heavy infestation reduce meat, milk and hide quality and overall host health. Although the precise monetary value of the economical impact is lacking, it is estimated that the global cost exceeds 10 billion US dollars annually (Sonenshine, 1991; Jongejan and Uilenberg, 2004; Bellgard *et al.*, 2012). Considering these facts, it is clear that it is of great importance to develop an effective tick control strategy. Vaccination is a method which offers several advantages over frequently used chemical acaricides (Bellés, 2010). Ultimately, the development of vaccines that target both the tick vector and the disease-causing pathogens may provide a means to simultaneously control tick infestation and pathogen transmission. However, one of the most limiting steps is the identification and characterisation of appropriate vaccine target antigens (de la Fuente and Kocan, 2006).

Tick proteins which are vital in physiological processes such as blood feeding and reproduction are potential targets for anti-tick vaccines, since vaccination could inhibit the vital functions of such proteins and ultimately affect tick survival (Bellés, 2010). This thesis reports the evaluation of the efficacy of metzincin metalloproteases (astacin and

reprolysin homologues) in the cattle tick, *Rhipicephalus microplus*, as possible anti-tick vaccine candidates.

1.2 Ticks in general

Ticks are classified into the class Arachnida, subclass Acari, superorder Parasitiformes and order Ixodida. Currently there are approximately 900 known species, which are subdivided into two major families, Ixodidae (hard ticks) and Argasidae (soft ticks) and a third smaller family, the Nuttalliellidae. The Ixodidae family is the largest, containing 13 genera and approximately 702 species. The Argasidae family comprises 5 genera and approximately 193 species (Moolhuijzen *et al.*, 2011). These two families can be differentiated according to several biological and behavioural criteria (Table 1.1). The Nuttalliellidae family consists out of a single species namely *Nuttalliella namaqua*, which is characterised by features mainly intermediate to those of the two major families (Barker and Murrell, 2004; Guglielmone *et al.*, 2010).

Table 1.1. The major differences between the two predominant tick families (Adapted from Sonenshine, 1991).

Ixodidae	Argasidae
1. Cuticle is relatively rigid.	1. Cuticle is smooth, highly folded and leathery.
2. Sclerotised dorsal scutum present.	2. Scutum absent.
3. Capitulum anterior and visible from dorsal view.	3. Capitulum subterminal, not readily visible from dorsal view.
4. Spiracles located behind fourth coxae.	4. Spiracles located between third coxae.
5. Larvae, nymphs and adults feed for several days.	5. Nymphs and adults feed rapidly (usually within minutes), while larvae can feed up to a few days.
6. Only one nymphal stage.	6. Two or more nymphal stages (up to eight in some species).
7. Males die after copulation. Females die after laying thousands of eggs (5 000 - 22 000 depending on species).	7. Several copulation acts. Females lay a few hundred eggs after each blood meal.

Due to the anatomical differences the feeding behaviours of hard and soft ticks differ quite drastically. Distinct from most other arthropods, ixodid ticks are slow feeders. While feeding, these ticks need to manufacture new cuticle to allow expansion for an adequate blood meal (Sonenshine, 1991). With argasid ticks the leathery, highly expandable cuticle allows rapid feeding (usually within minutes) and engorgement of

large volumes of blood, increasing the tick's size 5 -10 times their original pre-feeding size (Sonenshine, 1991).

The lifecycle of a tick constitutes four stages, the embryonated egg, followed by three active stages namely, the 6-legged larva, nymph and mature adult (Figure 1.1). In most species, each consecutive active lifestage is spent on a new host (multi or three-host lifecycle), the lifecycles of such ticks are slow and range from six months to several years. In a few species the larvae and nymphs remain and develop on the same individual host (one or two-host lifecycle), resulting in much more rapid lifecycle durations (Sonenshine, 1991). The latter is the case for *R. microplus.*

Ticks possess many remarkable features (physical and molecular), which contribute to their great success as ectoparasites. For successful feeding, a tick mechanically attaches to the host with the aid of mouthparts (chelicerae and hypostome). A feeding lesion is then created at the point of attachment, by penetrating the host's skin and

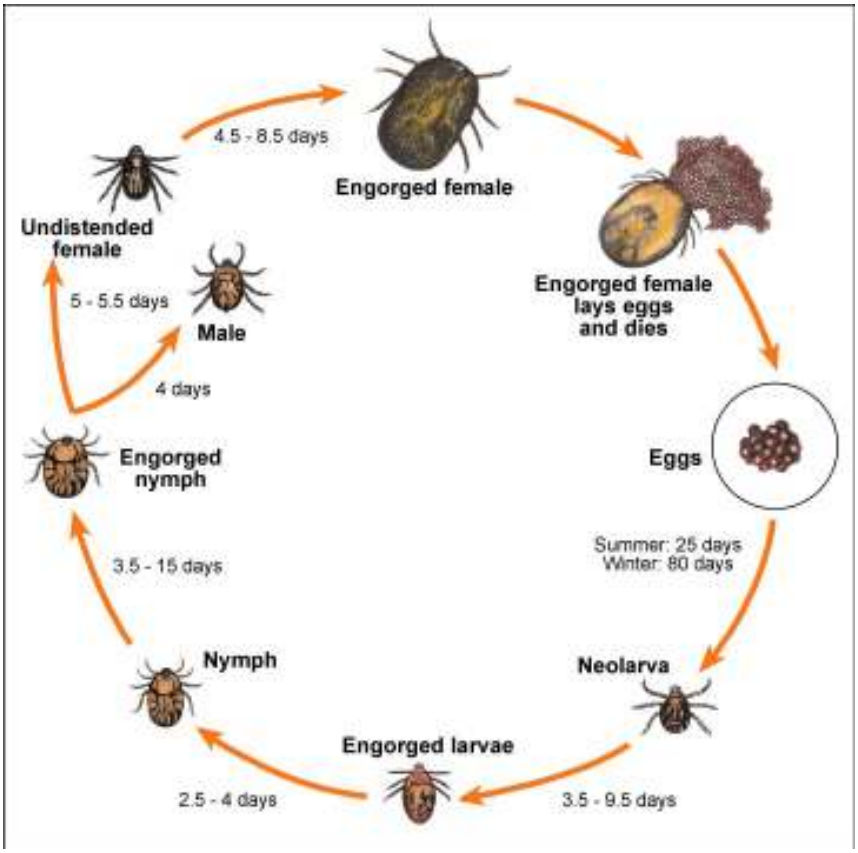

Figure 1.1. The four stage lifecycle, typical of the Ixodidae family. The indicated timescale is typical of *R. microplus*.

damaging blood vessels for the release of blood into a feeding pool. The latter results in the activation of the host's haemostatic system which includes blood coagulation and platelet aggregation. As a counter act, in order to maintain a fluid feeding cavity, ticks secrete several bioactive substances such as vasodilators, anti-coagulants, platelet aggregation inhibitors and fibrin(ogen)olytic agents (Table 1.2) (Maritz-Olivier *et al.*, 2007).

To assure sufficient nutrition during long starvation periods, ticks have attained an unusual yet remarkable digestion process. In contrast to haematophagous insects, ticks digest their blood meals almost entirely intracellularly and aside from egg laying females, it is relatively slow (Horn *et al.*, 2009). The blood meal is sucked in by a pharyngeal pump mechanism and pass through the oesophagus into the midgut lumen (Sonenshine, 1991). To allow rupture of the erythrocytes and subsequent release of haemoglobin, but still to prevent blood clot formation within the lumen, specialised secretory cells in the midgut epithilium secretes haemolysins and anti-coagulants, respectively. Haemoglobin and other proteins are endocytosed by type one digestive cells (D-I), which have highly modified plasma membranes with numerous microvilli and coated pits. The D-I cells mature into type two digestive cells (D-II), which are specialised cells capable of intracellular digestion and long-term storage of reserve nutrients for when required by the tick during starvation (Agbede and Kemp, 1986; El Shoura, 1988; Horn *et al.*, 2009).

Finally to maintain large tick populations, ticks have the remarkable ability to produce large numbers of eggs. To assure optimum fecundity, oogenesis is closely correlated with adult female feeding and in ixodid ticks only takes place post-engorgement (Sonenshine, 1991).

Table 1.2. Summary of the anti-haemostatic, vasodilator and immune-modulator components present in the saliva of ixodid ticks.c

	Common name	Source species	Mr (kDa)	References
Anti-Haemostatics	**Thrombin inhibitors**			
	Calcartin	*Boophilus calcaratus*	14.5	(Motoyashiki *et al.*, 2003)
	Boophilin (G2 and H2)	*R. microplus*	G2: 14 H2: 14	(Macedo-Ribeiro *et al.*, 2008)
	Amblin	*Amblyomma hebraeum*	17.4	(Lai *et al.*, 2004)
	FX, FXa and tissue factor pathway inhibitors (TFPI)			
	Ixolaris	*Ixodes scapularis*	15.7	(Francischetti *et al.*, 2002)
	Penthalaris	*I. scapularis*	35	(Francischetti *et al.*, 2004)
	Salp 14	*I. scapularis*	14	(Narasimhan *et al.*, 2002)
	Platelet aggregation/ adhesion inhibitors			
	Variabilin	*Dermacentor variabilis*	5	(Wang *et al.*, 1996)
	Apyrase	*I. scapularis*	n/d	(Ribeiro *et al.*, 1985)
	Longicornin	*Haemaphysalis longicornis*	16	(Cheng *et al.*, 1999)
	PGI2	*I. scapularis*		
	PGD2	*Amblyomma americanum*		
Vasodilators				
	Prostaglandin (PGE2)	*Ixodes dammini*	n/d	(Ribeiro *et al.*, 1985)
	Prostaglandin (PGE2 and PGF2α)	*A. americanum*	n/d	(Ribeiro *et al.*, 1992)
	Prostaglandin (PGE2)	*R. microplus*	n/d	(Dickinson *et al.*, 1976)
Immuno-modulators	**Histamine-binding proteins**			
	Male-specific RaHBP	*Rhipicephalus appendiculatus*	22.9	(Paesen *et al.*, 1999)
	Female-specific 1 RaHBP	*R. appendiculatus*	21.4	(Paesen *et al.*, 1999)
	Female-specific 2 RaHBP	*R. appendiculatus*	21.5	(Paesen *et al.*, 1999)
	25 kDa SG protein B	*I. scapularis*	25.5	(Das *et al.*, 2001)
	Host cytokine homologues			
	Tick MIF	*A. americanum*	12.6	(Jaworski *et al.*, 2001)
	DVHRF	*D. variabilis*	19.7	(Mulenga *et al.*, 2003)
	Immunoglobulin-binding proteins			
	IgG binding protein C	*R. appendiculatus*	19.3	(Wang *et al.*, 1996; Wang and Nuttall, 1999)

*Abbreviations correspond to: PGI/PGD, prostacyclin receptors; RaHBP, *R. appendiculatus* histamine-binding proteins; SG, salivary gland; MIF, macrophage migration inhibitor factor; DVHRF, *D. variabilis* histamine-release factor.

1.3 Rhipicephalus microplus

R. microplus (Figure 1.2), the tick of interest for this study, is an ixodid tick commonly known as the "cattle tick". This tick species causes annual economic losses in the hundreds of millions of dollars to cattle producers throughout the world, and ranks as the most economically important tick globally (Jongejan and Uilenberg, 2004; Carthew and Sontheimer, 2009).

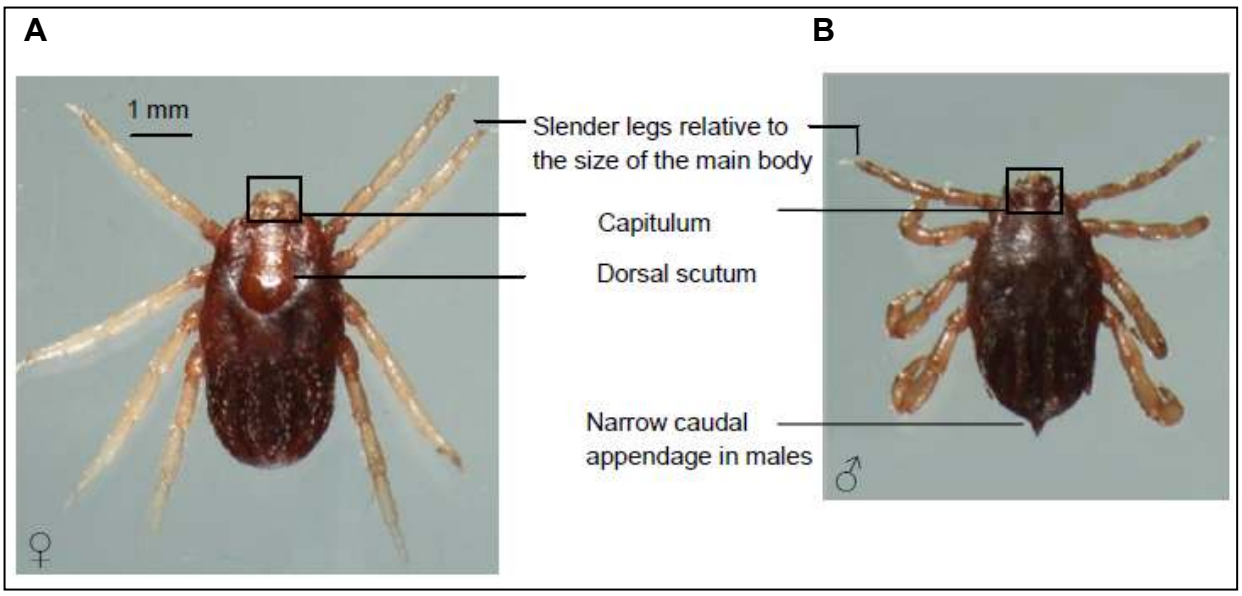

Figure 1.2. Photo-images of a freshly molted female (A) and male (B) *R. microplus* adult tick.

Historically *R. microplus* was introduced into North Australia from Indonesia and into East and South Africa from southern Asia via Madagascar, in the late 19th century after the 1896 rinderpest outbreak (Hoogstraal, 1956). Currently *R. microplus* is endemic in most subtropical and tropical regions of the world including: India, Asia, north eastern Australia, Madagascar, southern and central America and in south eastern Africa (Estrada-Peña *et al.*, 2006; Jose *et al.*, 2009) (Figure 1.3). Continued expansion of the distribution of *R. microplus* accompanied by the displacement of *Rhipicephalus decoloratus* has been reported over the past three decades in several African countries, including Zambia (Macleod and Mwanaumo, 1978; Berkvens *et al.*, 1998), Zimbabwe (Mason and Norval, 1980; Katsande *et al.*, 1996), Swaziland (Wedderburn *et al.*, 1991) and South Africa (Tønnesen *et al.*, 2004). Recent survey data indicates that *R. microplus* is still extending its distribution range and that high suitability is currently recorded for previously non-occupied areas, such as northern Tanzania (Lynen *et al.*, 2008) and West Africa (Madder *et al.*, 2011) (Figure 1.3). One external factor which

could have attributed to the tick's complicated ecosystem toward expansion might be the subtle changes in the climate, associated with global warming (White *et al.*, 2003).

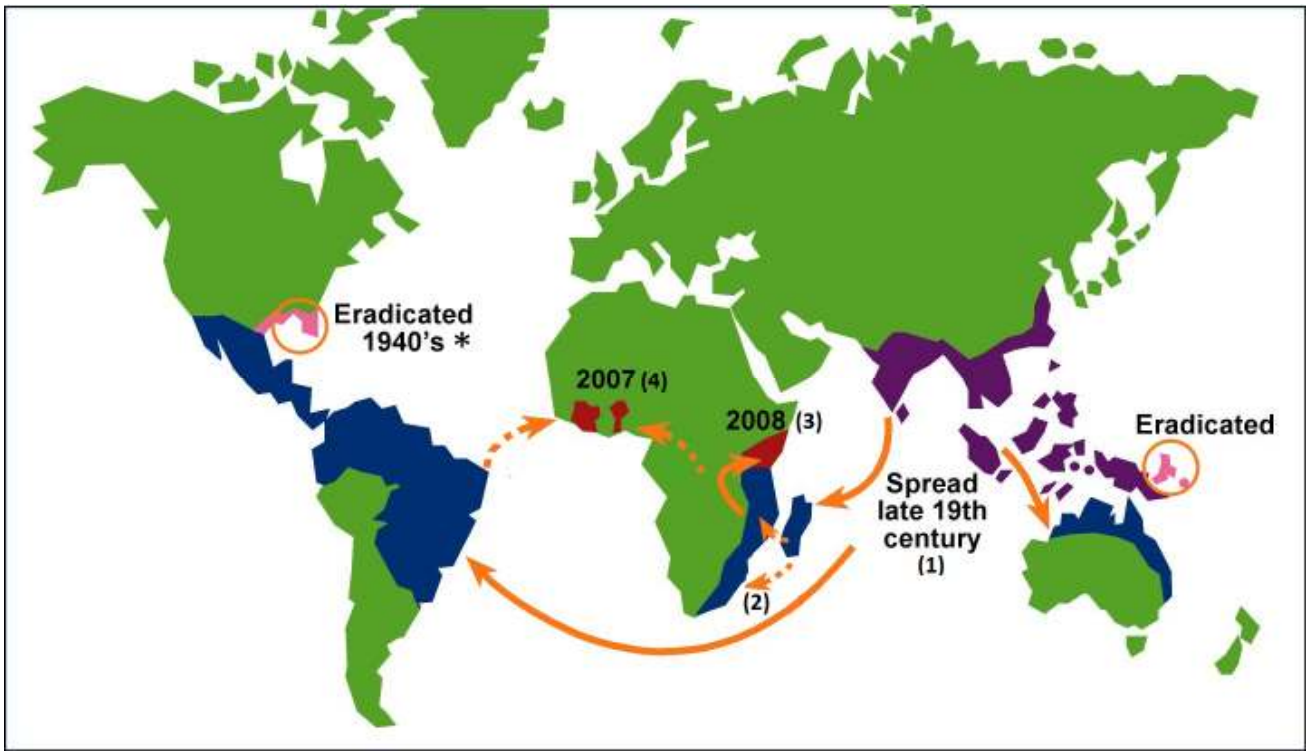

Figure 1.3. Geographical distribution of *R. microplus*. A number of references were used to map the movement of *R. microplus* spreading: (1) Hoogstraal (1956), (2) Tønnesen *et al.* (2004), (3) Lynen *et al.* (2008) and (4) Madder *et al.* (2007). Dashed arrows depict that the precise routes of spread are unknown. * In 2009 a temporary expansion of the quarantine zones have been required due to reinfestations caused by the spread of *R. microplus* by white-tailed deer populations (Pound *et al.*, 2010).

R. microplus is an ectoparasite of great concern, owing to the fact that it serves as vector of important livestock disease pathogens including, *Anaplasma marginale, Borrelia theileri* and *Babesia spp.* (Jongejan and Uilenberg, 2004). While *R. decoloratus* only efficiently transmits *Babesia bigemina, R. microplus* is a competent vector of both *B. bigemina* and *B. bovis* and since the latter protozoan is more pathogenic and in effect results in much heavier losses in susceptible livestock, the great concern regarding *R. microplus* displacing *R. decoloratus* is obvious.

Bovine babesiosis (also known as cattle tick fever, Texas fever, piroplasmosis or redwater) once devastated the USA cattle industry (Friedhoff and Smith, 1981) and still remains the most economically relevant arthropod-borne disease of cattle worldwide, due to significant morbidity and mortality (Bock *et al.*, 2004; Uilenberg, 2006; Calixto *et al.*, 2010). This infection is characterised by high fever, ataxia, anorexia and general

circulatory shock. *B. bovis* infection can also lead to central nervous system signs, as a result of sequestration of infected erythrocytes in the brain capillaries, which usually suggest poor prognosis. In the last decade, human babesiosis has been identified as a rapidly emerging zoonotic infectious disease – causing malaria-like symptoms in humans (Kjemtrup and Conrad, 2000). Several clinical syndromes have been described, ranging from asymptomatic infections to persistent severe relapsing babesiosis manifesting mostly in individuals with malignancy, HIV or immuno-compromised patients. Although the rodent strain *B. microti* is the most common species responsible for human babesiosis, the severity of infection is variable. However, infection with the cattle strains *B. divergens and B. divergens*-like parasites have been detected and have shown to cause a severe form of babesiosis, ending in death in more than a third of patients (Kjemtrup and Conrad, 2000). The emergence of human babesiosis from an increasing number of *Babesia spp.* reinforces the great concern of the expanding distribution range of the tick vectors.

Besides the pathogens *R. microplus* transmits, heavy tick loads also cause direct harm to hosts as their feeding behaviour results in blood loss, scar tissue formation and leather damage. As *R. microplus* is a one-host tick each active lifestage feeds only once and therefore the whole lifecycle can be completed in as short as 3-4 weeks, resulting in heavy seasonal tick burdens. Currently, control of *R. microplus* in general relies on chemical acaricides. However, over the past three decades *R. microplus* has developed resistance to all major classes of acaricides (Guerrero *et al.*, 2012a) including, organophosphates (Li *et al.*, 2003), pyrethroids (Beugnet and Chardonnet, 1995; Miller *et al.*, 1999) and formamidines (Kunz and Kemp, 1994; Li *et al.*, 2004). This resistance could be attributed to the intensive use of these chemical substances and/or be due to the rapid lifecycle of *R. microplus*, given that the selection pressure is directed against all life stages (Peter *et al.*, 2005). Different acaricide resistance strategies have been considered including rotation of acaricides and mixtures of acaricides or synergists (Li *et al.*, 2007). However, several *R. microplus* populations have already been shown to develop cross or multiple resistance, due to their biological and behavioural characteristics (Baffi *et al.*, 2007). Cross-resistance results from resistance to two or more classes of acaricides with a similar mode of action, while multi-resistance is due to resistance to two or more classes of acaricides with differential action mechanisms (Rodríguez *et al.*, 2002; Yoon *et al.*, 2004).

Considering all of the above, it is clear that it is of great importance to develop an alternative effective control strategy, such as an appropriate vaccine against these ticks.

1.4 Tick control

The control of tick infestation poses a significant challenge for the cattle industry, since the natural enemies of ticks (which include: birds, rodents, shrews, ants and spiders) do not possess the ability to significantly reduce tick populations. To date, tick control has mainly been achieved through the use of acaricides. In the short term this option is often the most convenient, economical and does not require a refined knowledge of the biology of the pest by the user. In the first half of the 20th century the main insecticides were arsenic based. Subsequently, organochlorides, organophosphates, carbamates, amidines, pyrethroid and avermectins made their way into the market and are still being used worldwide today (Graf *et al.*, 2004). In many countries the sale of these chemical substances accounts for almost two thirds of the veterinary market (Peter *et al.*, 2005). However, this control measure has had limited efficacy and holds several disadvantages.The first is that routine application of acaricides is labor intensive. This is followed by the delayed degradation of chemical residues that therefore remain for long periods in the agricultural environment, where it adversely affects meat and milk products as well as the water and life of living organisms in the natural ecosystem. Finally, the sustainability of this approach is diminished by the resistance ticks develop against acaricides.

In recent years, survey data have indicated that globally tick populations have developed resistance to many of the different classes of acaricides (Rodríguez *et al.*, 2002; Li *et al.*, 2003; Li *et al.*, 2004; Rosario-Cruz *et al.*, 2009; Klafke *et al.*, 2012). Target site mutations are the most common resistance mechanism observed, but resistance can also arise through metabolic mechanisms or penetration resistance (Guerrero *et al.*, 2012a). Although the mechanism of penetration resistance has been identified in artropods, including *R. microplus,* the excact mechanism remains unknown. Resistance most likely result by means of alterations in the ability of an acaricide to penetrate or enter the treated individual organism (Rosario-Cruz *et al.*, 2009). Target site resistance occurs when an allele of the gene coding for the target molecule has an amino acid mutation, confering resistance to the acaracide. This mechanism of resistance is particularly well studied in the case of pyrethroid-based acaricides, where

the voltage-gated sodium channel is the target site (Morgan *et al.*, 2009). Metabolic resistance takes place when the ability of an individual to detoxify or sequester an acaricide is altered (Guerrero *et al.*, 2012a). Enzymes known to be involved in metabolic resistance include esterases, glutathione S-transferase and cytochrome P450s. The main resistance mechanisms in *R. microplus* include point mutations in esterase-encoding genes (Hernandez *et al.*, 2000; Jamroz *et al.*, 2000; Villarino *et al.*, 2003), octopamine receptor (Chen *et al.*, 2007), para-sodium channel (Morgan *et al.*, 2009) and gama-aminobutyric acid (GABA)-receptor (Hope *et al.*, 2010).

The increasing worldwide resistance to acaricides therefore necessitates greater research on the identification of new improved acaricide targets. But, the ultimate drawback of this control approach is that the development of new acaricides is a long and very expensive process (Graf *et al.*, 2004). Thus, research on alternative and safer strategies which could be used in an integrated manner is crucial for controlling tick infestation (Willadsen, 2006).

Alternative approaches include pasture rotation and the use of hosts with natural resistance to ticks. But, even though pasture rotation has been used effectively for the control of *R. microplus* in Australia (Jonsson, 1997), the practicability of this method is limited in third world countries in view of the fact that it requires very high initial start-up costs and active management of livestock. In areas where ticks and tick-borne diseases are endemic and control relies on the use of hosts with natural-acquired resistance (e.g. *Bos indicus*), importation of improved cattle breeds (but with high susceptibility, e.g. *Bos taurus*), renders this approach to be ineffective (Willadsen, 2006). Of interest is research leading to the development of novel tick-control strategies which are based on the use of biological control agents, such as entomopathogenic fungi of the genera *Beauveria* and *Metarhizium* (Frazzon *et al.*, 2000; Benjamin *et al.*, 2002; Gindin *et al.*, 2002). Although biocontrol strategies are in principle highly desirable, the manufacture, efficacy, application and stability of such agents in a field situation presents serious challenges for the practical use of such a control method (Benjamin *et al.*, 2002).

One of the most attractive alternative control approaches is to exploit the host's immunity (Willadsen, 2004). The first example of experimentally immunising animals against ticks dates back to 1939, where Trager and colleagues showed that crude tick tissue extracts could be used to immunise guinea pigs against *Dermacentor*

variabilis (Trager, 1939). Fifty years later it was demonstrated that the same principle can be applied to protect cattle against *R. microplus* (Johnston *et al.*, 1986). Although promosing, the method of injecting whole crude extracts from *R. microplus* soon proved to yield only moderate succes (Elvin and Kemp, 1994). However this groundbreaking work proved that an anti-body mediated protective response takes place, supporting the hypothesis that a vaccine against *R. microplus* could be developed. Control of ticks by vaccination has the advantages of being cost-effective, reducing environmental contamination and preventing the selection of drug-resistant ticks that result from repeated acaricide application. Furthermore, by reducing tick populations and/or affecting tick vector capacity, vaccines can also indirectly prevent or reduce transmission of pathogens (de la Fuente *et al.*, 1998; de la Fuente *et al.*, 2007a).

De Rose *et al.* (1999) demonstrated the feasibility of controlling tick infestation by immunisation of the host with a selected tick gut antigen. To date, the only anti-tick vaccine (commercialised as TickGARD™ in Australia and GAVAC™ in South America) is a recombinant vaccine based on the midgut protein Bm86 from *R. microplus.* Although the vaccine was not sufficient to eradicate ticks, it reduced subsequent generations by significantly affecting fecundity (Willadsen, 2004; Willadsen, 2006). However, the variable efficacy of Bm86-based vaccines against different geographic *R. microplus* strains (García-García *et al.*, 1999), together with the need to develop an effective combined anti-tick and anti-pathogen vaccine have encouraged the research for additional tick protective antigens. The goal is not necessaraly total eradication by a single vaccine, but rather that vaccination together with integrated approaches will accomplish endemic stability of both the tick vector and diseases causing pathogens.

The most limiting step for the improvement and/or development of new vaccines, however, remains the identification and characterisation of effective protective antigen targets. Tick antigens studied to date are from a limited range of functional classes, including structural proteins (particularly of the salivary glands), various proteases and their inhibitors (particularly those involved in haemostatic processes) and a range of membrane-associated proteins of unknown function (Maritz-Olivier *et al.*, 2007; Bellés, 2010). Potential targets which have been investigated in recent years include, regulatory proteins, such as ubiquitin and transcription regulators (e.g. subolesin and elongation factor-1 alpha) (Almazán *et al.*, 2010). By regulating gene expression, these proteins are indirectly involved in the control of multiple cellular pathways and

processes regulating tick blood digestion, reproduction and development. Other novel targets that are worth investigating comprise tick proteins that directly act in processes such as blood feeding, digestion, reproduction and development, in view of the fact that vaccination could inhibit these essential functions and ultimately affect tick survival (Bellés, 2010). An example inculde proteases, including metzincins, that will be discussed below.

1.5 Metzincins, a clan of metalloproteases

Proteases (also termed peptidases, proteinases and proteolytic enzymes) are enzymes which irreversibly hydrolyse peptide bonds of polypeptides. These proteolytic enzymes were initially thought to be simple destructive enzymes necessary only for non-specific degradation in processes such as digestion and intracellular protein turnover. However ground breaking work in the middle of the last century (Davie and Neurath, 1955), uncovered that proteases also catalyse highly specific reactions. Multiple roles of regulated proteolysis include the regulation of protein localisation and activity; modulation of protein-protein interactions; generation and amplification of intracellular signals and the modulation of the extracellular environment (Neurath and Walsh, 1976). By either highly specific or non-specific polypeptide cleavage, proteases play key roles in virtually all important physiological processes, such as cell growth (cell proliferation and differentiation), tissue development (angiogenesis, neurogenesis and tumorigenesis), reproduction (ovulation and fertilisation), wound healing (haemostasis), immunity and apoptosis (Barrett *et al.*, 2004).

Proteases are either exopeptidases cleaving the protein substrate from the N-terminus (aminopeptidase) or C-terminus (carboxypeptidase), or endopeptidases that cleave the protein substrate internally (Figure 1.4).

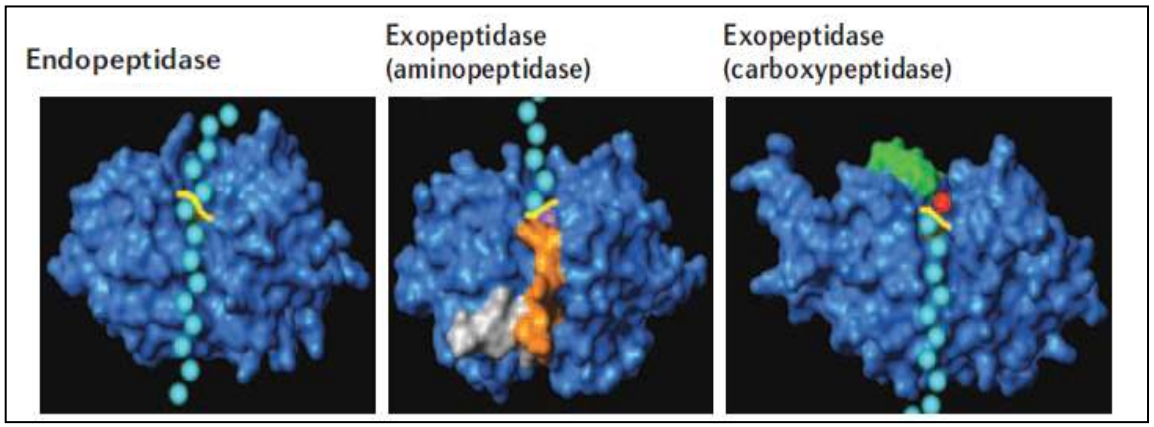

Figure 1.4. Space-fill models of substrate cleavage by proteases (Turk, 2006). Peptide substrate (schematically represented by cyan balls) runs through the entire length of the active site of an endopeptidase framework (blue) and is cleaved in the middle of the molecule (scissile bond marked yellow). In exopeptidases, substrate binding is constrained by structure elements (represented in orange and green), which also provide negative charge to bind to positively charged amino-terminus (purple) of the substrate, or positive charge to bind the negatively charged carboxyl terminus (red) of the substrate.

Based on the mechanism of catalysis, proteases are classified into six classes: the aspartic-, cysteine-, serine-, threonine-, metallo- and glutamic proteases (the glutamic proteases being the only not yet to be described in mammals) (López-Otín and Bond, 2008). The proteases of each class are further grouped into families and clans. Proteases which have significantly similar amino acid sequences around their active sites are grouped into a family. Based on their three-dimensional structures, one or more families can be grouped into a clan (Barrett *et al.*, 2004).

Bioinformatic analysis of the mouse and human genomes has revealed that about 2% of all gene products are proteases, of which approximately one third are metalloproteases (Barrett *et al.*, 2004; Rawlings *et al.*, 2008). Metalloproteases (EC.3.4.24) are enzymes which require the presence of a metal ion for full catalytic activity, usually zinc but in some cases also cobalt, manganese, nickel or copper (Jongeneel *et al.*, 1989). The majority of zinc-dependent metalloproteases (so-called zincins) contain the short zinc-binding consensus sequence, HEXXH (X: any amino acid), in their active sites (Bode *et al.*, 1993). Only a small group of zinc-dependent metalloproteases (the inverzincins) are characterised by the inverted zinc-binding motif, HXXEH (Figure 1.5) (Hooper, 1994).

The different families of the zincins have been grouped according to the nature and the position of their third zinc-binding residue into three clans: the metzincins (MEROPS: mainly the MA(M) subclan), gluzincins (MEROPS: MA(E)) and aspzincins (MEROPS:

part of MA(M) subclan) (Gomis-Rüth, 2003). The largest of three, the metzincins is the clan of interest for this study.

Metzincins are multi-domain endopeptidases, which are distinguished by an extended zinc-binding consensus motif (HEXXHXXG/NXXH/D) in the active site and a conserved methionine-containing turn (Met-turn) underlying the active site (Bode *et al.*, 1993; Gomis-Rüth, 2009). This clan consists of several zinc-dependent metalloprotease families, which exhibit significant topological similarity around the catalytic site, creating virtually identical zinc-binding environments (Stöcker and Bode, 1995). To date, seven metzincin families have been analysed at the structural level, including astacins, reprolysins, serralysins, matrix metalloproteases (MMPs), leishmanolysins, snapalysins and pappalysins (Gomis-Rüth, 2009) (Figure 1.5).

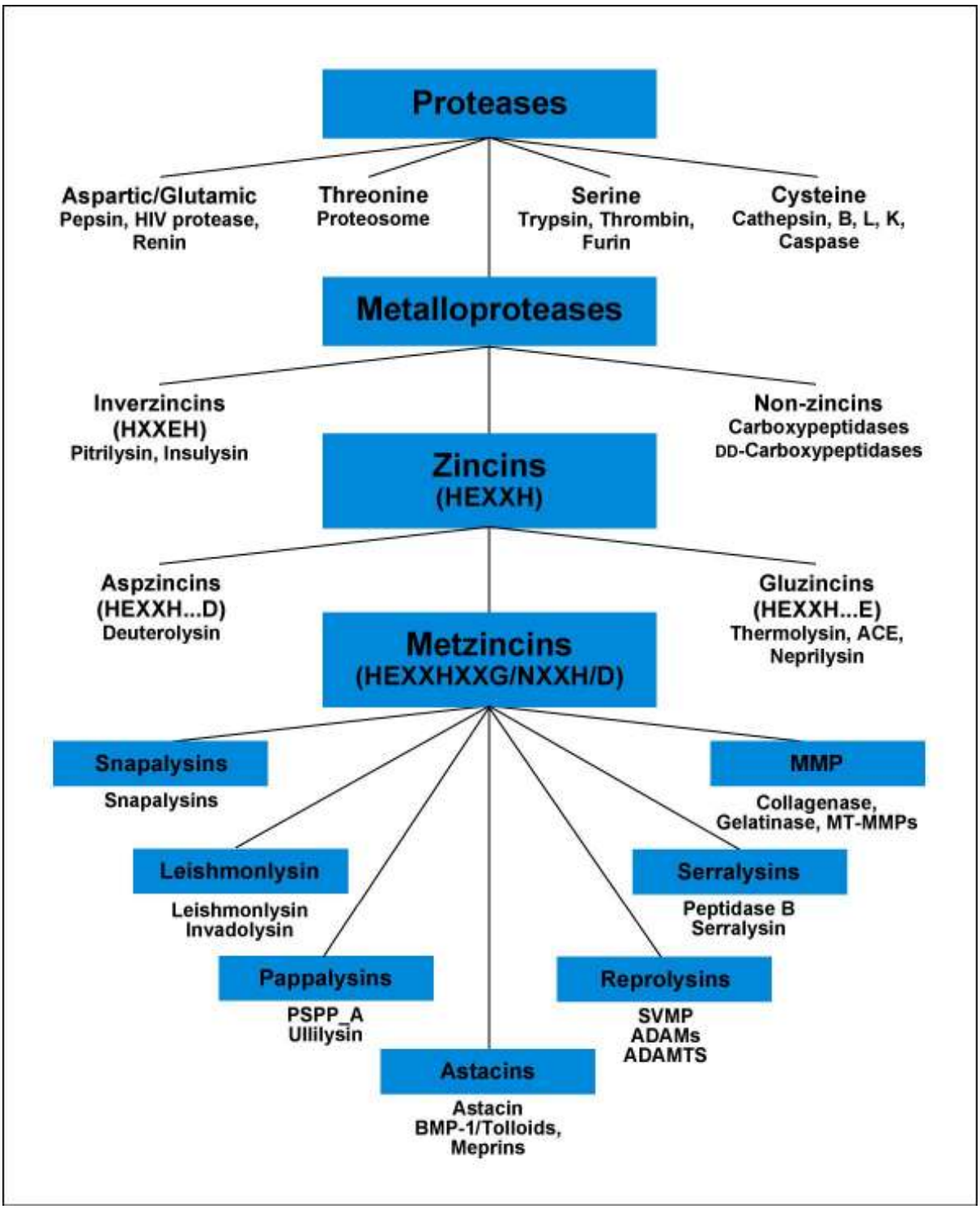

Figure 1.5. Classification scheme of proteolytic enzymes, highlighting the metzincin clan (Adapted from Sterchi *et al.*, 2009). Abbreviations correspond to: X, any amino acid; ACE, angiotensin-converting enzyme; MMP, matrix metalloprotease; MT, membrane bound MMP; SVMP, snake venom metalloproteases; ADAMs, a disintegrin and metalloproteases; ADAMTS, ADAMs containing trombospondin-like domains; BMP-1, Bone morphogenetic protein-1; PSPP_A, pregnancy-associated plasma protein A.

Members of the metzincin clan have been identified in a wide variety of living organisms including plants, fungi, bacteria and animal species, ranging from protozoa to humans. In the different biological systems, these metalloproteases play key roles in vital physiological processes such as tissue and neuro development, feeding and digestion, embryo development and egg hatching, and proteolytic cleavage of the extracellular matrix (ECM) (Bowles *et al.*, 2008; Sterchi *et al.*, 2009; Aung *et al.*, 2011).

Coherent with these essential roles, failure of regulation of these enzymes leads to multiple pathological conditions like cardiovascular and inflammatory diseases, neurodegenerative disorders and cancer (McQuibban *et al.*, 2000; Chang and Werb, 2001; Blobel, 2005; Chernov *et al.*, 2010; Endres and Fahrenholz, 2010). Consequently, metzincins are recognised as attractive drug targets for these diseases, and are also considered as potential diagnostic and prognostic biomarkers, especially in cancer (Turk, 2006). Furthermore, since metzincins have been recognised as key players in reproduction and host infestation of harmful parasites and because many disease causing micro-organisms require metzincins for replication and use them as virulence factors, these proteases have been considered targets for control strategies against such organisms. The applicability of metzincins as vaccine candidates, against these medically and economically important organisms, will be discussed in Section 1.6.

1.5.1 Structural features of metzincin catalytic domains

1.5.1.1 Overall topology

Since the first metzincin structure, that of astacin (EC 3.4.24.21), has been solved to 1.8 Å resolution (Bode *et al.*, 1992), the interest in metzincins considerably increased. Up to date more than 100 structures, comprising the characteristic metzincin catalytic domain, have been submitted to the Protein Data Bank (http://www.rcsb.org/pdb). Comparative analysis revealed that despite a low overall sequence similarity (< 30%), proteases of the metzincin clan share a common scaffold and well conserved active site environment (Stöcker *et al.*, 1995). However each family has distinguished structural elements, of which the specific features will be discussed in Chapter 2.

The overall shape of a metzincin protease comprises two domains, the upper N-terminal subdomain and the lower C-terminal subdomain, which is separated by the central active site cleft. The N-terminal subdomain consists of a twisted β-sheet and two long α-helixes (Figure 1.6). The sheet typically consists of five β-strands, except in leishmanolysins, which only have four (Gomis-Rüth, 2003). β-I, II, III and V are parallelly orientated to one another and bind substrate. β-II is positioned the furthest from the active site cleft and exhibits the largest variability of the regular secondary structure elements in both length and spatial orientation. β-IV is orientated anti-parallel to the other strands and forms the lower border of the N-terminal subdomain, creating the upper edge (northern wall) of the active site cleft, and binds the substrate, mainly on its non-primed side, in an anti-parallel manner (Figure 1.6) (Stöcker *et al.*, 1995). The greatest variability among the different metzincin families occurs in the segment connecting β-III and β-IV. This loop has unique bulge-like elements for each family, that affect substrate binding in the S_1' and S_2' pockets and consequently affect the specificity of the proteases (Stöcker *et al.*, 1995; Gomis-Rüth, 2009).

The two α-helices are arranged on the concave side of the β-sheet, in an identical manner throughout all metzincins, and are almost completely superimposible. Helix αA connects β-II and β-III, resulting in a right handed β-α-β motif. Helix αB supplies two of the zinc-liganding histidines and separates them from one another by a single helix turn, positioning the imidazole side chains toward the catalytic zinc (Stöcker *et al.*, 1995). After the active site helix αB the main chain sharply turns down into the C-terminal subdomain at a strictly conserved glycine residue. In such a low-energy conformation the main-chain angles of the glycine residue allows this sudden turn and explains its conservation (Figure 1.7).

The C-terminal subdomain continues, after the conserved glycine residue, with a loop that contains the third zinc-liganding histidine. This loop, which length can vary from 6 to 53 residues, ends near the catalytic zinc-ion in a unique 1, 4-turn. This right-handed screw turn contains a methionine residue, which is strictly conserved and therefore has been termed the “Met-turn” (Bode *et al.*, 1993). This unique turn is a structural signature for this clan and gave rise to the name, the “metzincins” (Bode *et al.*, 1993). The rest of the C-terminal subdomain contains few regular secondary-structure elements and varies considerably in size. The only conserved element is a long α-helix (αC) situated close to the C-terminal (Figure 1.6).

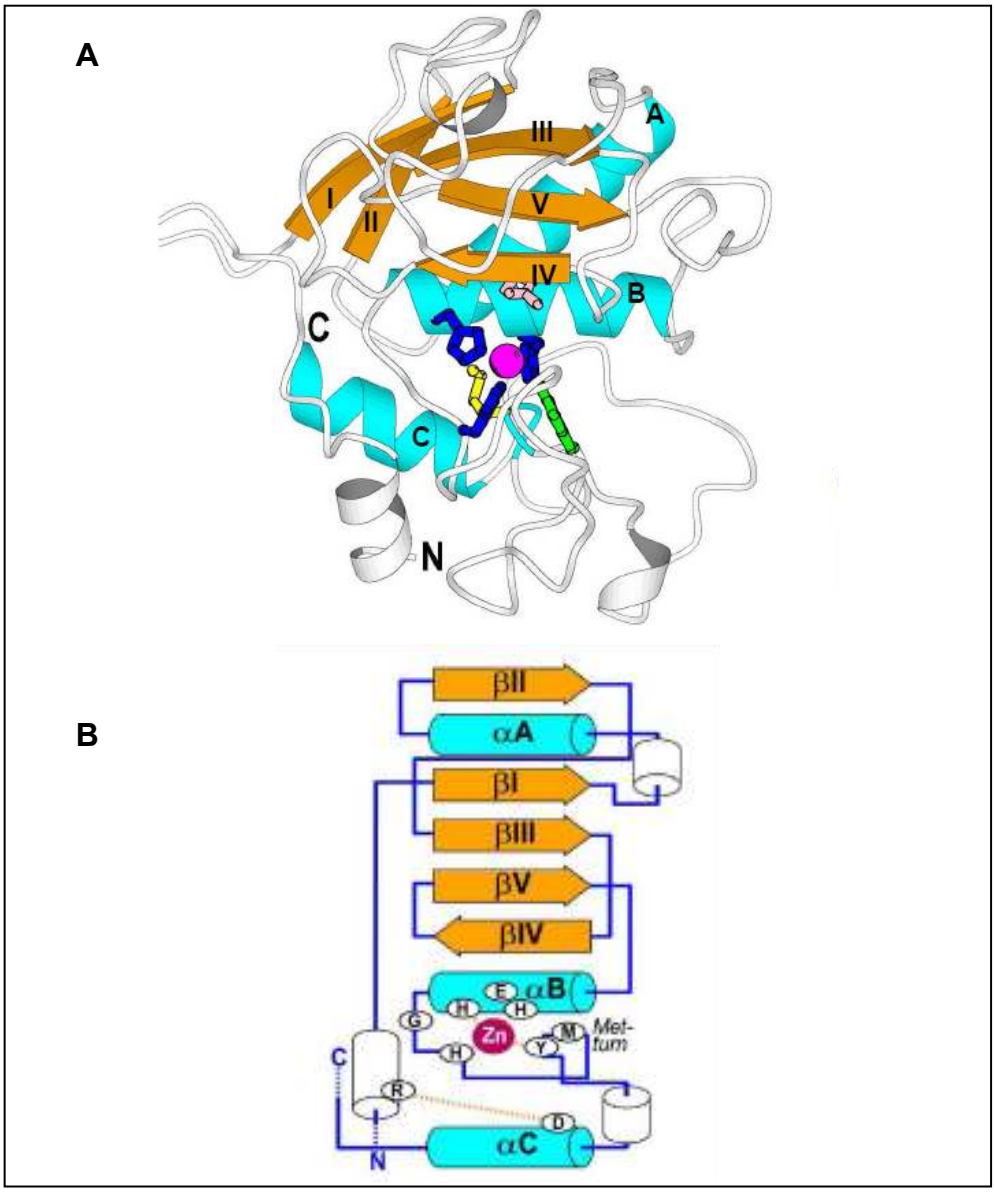

Figure 1.6. Richardson diagram (A) and topology scheme of aeruginolysin (PDB 1kap) (B), presenting the characteristic metzincin catalytic-domain (Adapted from Gomis-Rüth, 2009). In figure A: the common β-strands and α-helices are shown in orange and cyan, respectively; the N- and C-termini are labeled in large capital letters; and the zinc-binding histidines (blue), general base (light pink), zinc-ion (magenta), zinc-binding tyrosine (green), Met-turn methionine and disulfide-bonding cysteines (yellow) are displayed. In figure B: helices are indicated as rods, strands as arrows, and amino acids as ellipses; common elements are depicted in cyan and orange as in (A) and are labeled (α-helices, from A to C, β-strands, from I to V, amino acids as dark ellipses with white lettering); disulfide bonds are shown as orange connections; and the distinguishing regular secondary structure elements of aeruginolysin are shown in white.

1.5.1.2 The zinc-binding site

The highest degree of primary structure similarity among the metzincins, is found in the active site region (Figure 1.7). In most metzincin families the active site zinc-ion is ligated by the Nε2 atoms of the three consensus histidine residues, while in the snapalysins the zinc-ion is coordinated with two histidine residues and one aspartate (Oδ2 atom). The distance between the zinc-ion and the imidazole side chains, of the different histidines, varies between 2.0 – 2.3 Å for different metzincin proteins (Gomis-Rüth *et al.*, 1993b). In unbound metzincins a catalytic water molecule, bound to a conserved glutamate, forms a fourth ligand to the zinc-ion. In most of the families, these four ligands distort the metal-ion in a trigonal pyramidal coordination sphere. In unbound astacins and serralysins, the Oη atom of a conserved tyrosine residue (within the met-turn) serves as a fifth ligand to the zinc-ion, expanding the coordination sphere to trigonal-bipyramidal geometry (Gomis-Rüth *et al.*, 1993b).

It is reasoned that the aromatic ring of the tyrosine plays a role in substrate binding and/or stabilization of the tetrahedral intermediate product. In metzincins lacking this tyrosine residue, a superimposable proline is found, that contributes to the stabilisation of the active site, but in this case it does not extend into the zinc-coordination sphere. The last residue of the elongated zinc-motif (following the third histidine) is known as the 'family-specific conserved residue' (usually indicated as Z), since it is strictly conserved within each subfamily (Glu in astacins, Asp in reprolysins, Ser in MMP and Pro in serralysins). These residues also contribute to the stabilisation of the active site (Figure 1.7) (Stöcker *et al.*, 1995).

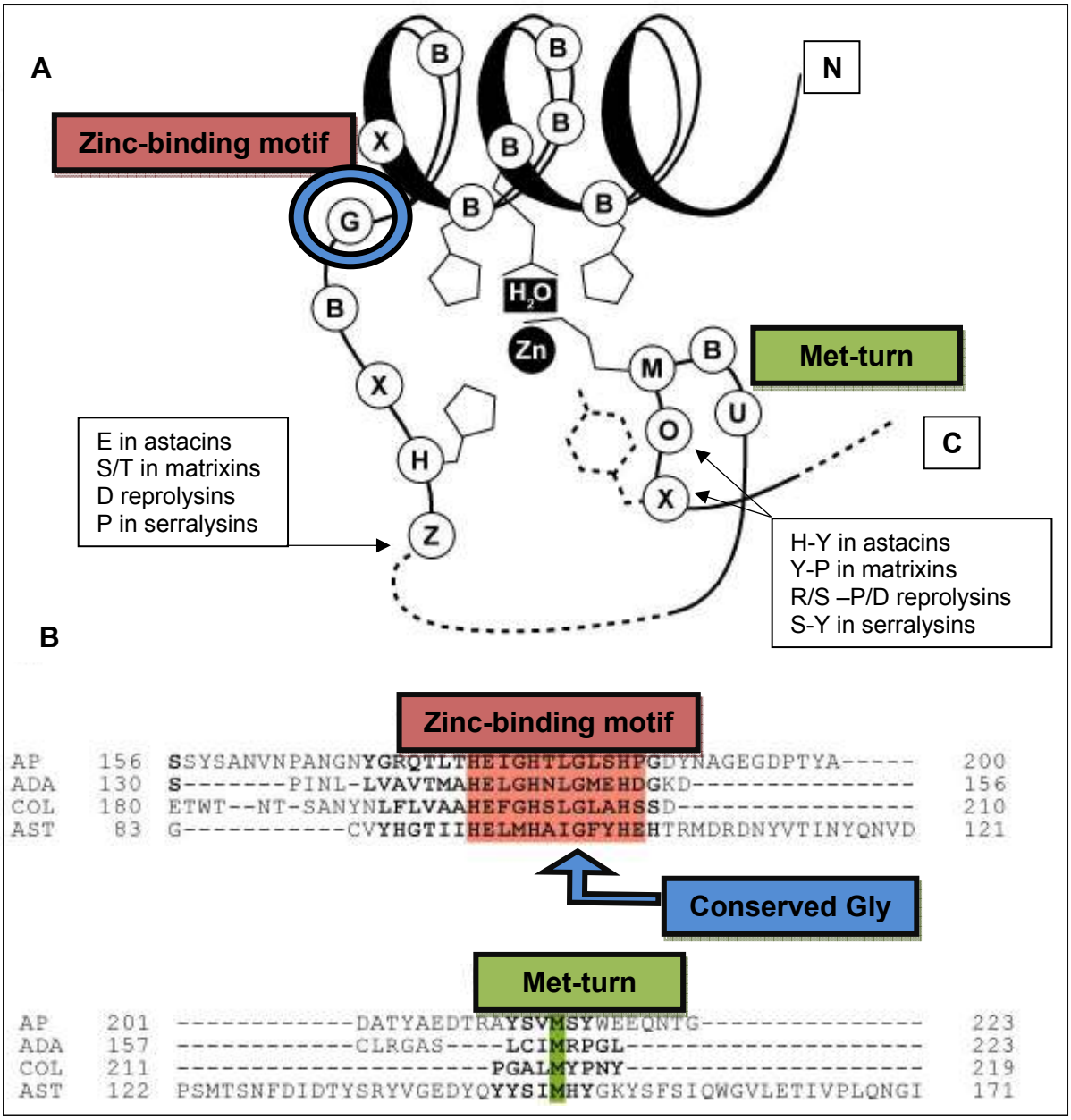

Figure 1.7. Scheme illustrating the zinc-binding environment and Met-turn of the metzincins (A) and a sequence alignment of the catalytic domains of four metzincins (Adapted from Stöcker and Bode, 1995). In figure A the abbreviations correspond to: B, bulky residues; X, any amino acid; Z, family-specific conserved residue. In figure B sequences correspond to: AP, alkaline proteinase from *Pseudomonas aeruginosa* (gi|416632); ADA, adamalysin II from *Crotalus adamanteus* (gi|547144); COL, collagenase from human neutrophil granulocytes (gi|180618); AST, astacin from *Astacus astacus* (gi|1200203|). The position of the elongated zinc-binding motif (pink), the Met-turn (green) and the conserved glycine (blue arrow) are indicated.

1.5.2 Substrate binding and catalytic mechanism of metzincins

Almost all proteases bind their substrates in a similar manner. The general nomenclature of cleavage site positions of the substrate and the binding pockets of the enzyme was formulated by Schechter and Berge in 1967. They designated the cleavage site between the P_1 - P_1' substrate residues and labeled the series of binding site S_x and S_x'. The non-primed S_x sites bind up stream of the scissile bond to the P_x peptide residues, incrementing the number in the N-terminal direction. The primed S_x' sites bind on the carboxyl side of the scissile bond to the P_x' peptide residue and the numbering are incremented in the C-terminal direction (Figure 1.8 A). An example is shown in Figure 1.8 B.

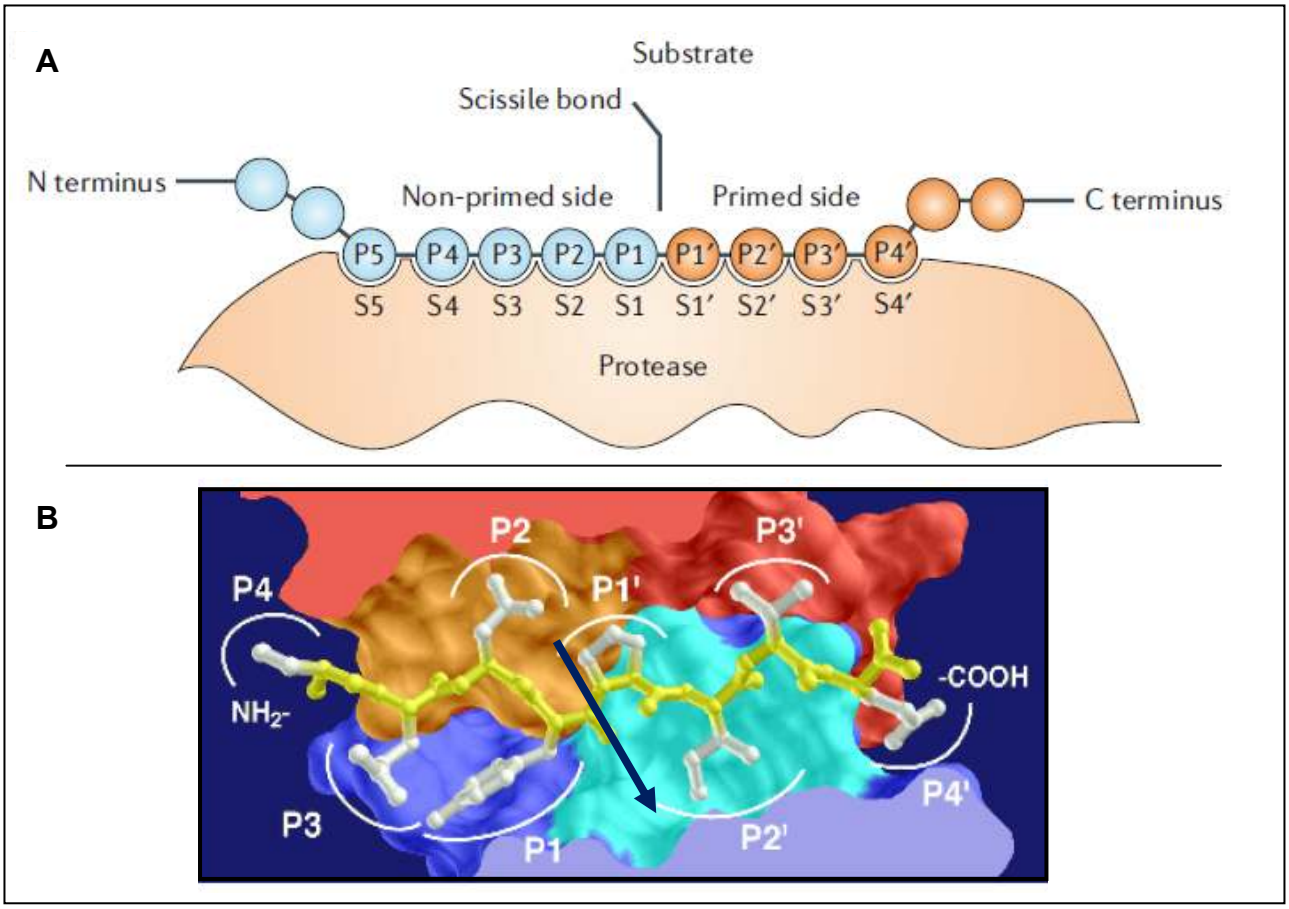

Figure 1.8. Schematic representation of a protein substrate binding to a protease (A) and as an example, the stereo view of the interaction between HIV-1 protease and a polypeptide substrate (B) (Adapted from Turk (2006) and www.scripps.edu/~stoffler/proj/Coev/CoevFig1.gif). In figure A, the subsites are numbered S_1–S_n upwards towards the N terminus of the substrate, and S_1′–S_n′ towards the C terminus, beginning from the sites on each side of the scissile bond. The substrate residues they accommodate are numbered P_1–P_n, and P_1′–P_n′, respectively. In figure B, the substrate binding pocket positions, and the designated cleavage site (navy arrow) are indicated.

The active site of metzincins is designed to accommodate substrates for subsequent hydrolysis. Crystal structure analysis of metzincins in complex with substrate analogs and peptide-based inhibitors has revealed that these enzymes bind their substrates horizontally to the active site cleft in an extended manner (Bode *et al.*, 1994; Grams *et al.*, 1995).

The enzyme-substrate interaction on the N-terminal side of the scissile bond is relatively similar amongst various metzincins. The non-primed side of the peptide substrate gets positioned on the bottom of the cleft, on top of the second zinc-liganding histidine, and aligns anti-parallel to the northern wall strand, β-IV of the active site cleft. The central four residues of the β-IV strand are relatively conserved (G/A/C-X-A/S-Y/F) and interact with the non-primed side of the substrate by means of hydrogen bonds, leaving the construct in a slightly twisted manner. Since the P_1 residue of each individual metzincin family interacts differently with the active site cleft walls, the S_1 substrate-binding pocket is relatively unique to each family. The S_2 subsite is represented by a relatively shallow surface depression of similar size and shape and the S_3 subsite is formed by a hydrophobic groove in the upper wall, which is particularly well suited to harbor a proline as the P_3 residue (Stöcker *et al.*, 1995). The C-terminal side of substrates appears to be bound in different ways amongst different metzincins. For adamalysin and some MMPs, the primed side of the peptide substrate is fixed between a bulge-like segment of the anti-parallel β-IV strand and the parallel lower wall (southern wall) by means of two hydrogen bonds. In astacins, side-chain interactions in the S_2' and S_3' subsites (and possibly also the fifth zinc-liganding tyrosine) are responsible for the primed site fixation (Grams *et al.*, 1995; Yiallouros *et al.*, 1998). Since all metzincins present multiple attachment sites along the substrate-binding region, these enzymes require extended substrates for optimum cleavage efficiency. Alignment and substrate binding takes place in order to bind the substrate in such a manner that the carbonyl group of the scissile bond is projected toward the catalytic zinc-ion to be polarised, for optimal substrate cleavage to take place.

In the native enzyme state a water molecule, designated as the activated solvent (Vallee and Auld, 1990) serves as the fourth zinc ligand and is also bound to the glutamate of the consensus zinc motif. Mutation studies revealed that the glutamate is essential for substrate cleavage, since mutation of this residue caused complete loss of proteolytic activity (Becker and Roth, 1992). Upon substrate binding the zinc-liganding

water molecule becomes squeezed between the carbonyl group of the substrate's scissile bond and the catalytic glutamate. The glutamate polarises the water molecule, which then serves as a nucleophile, attacking the carbonyl of the scissile bond via its lone pair electrons (Figure 1.9). The zinc-ion acts as a Lewis base by accepting the pair of electrons, thereby becoming temporarily penta-coordinated. This transition state is stabilised by the positive charge of the ion and nearby amino acid residues, such as the conserved tyrosine in astacins and serralysins. Finally, glutamate acts as a general base and shuttles protons to the scissile amide nitrogen atom to break the peptide bond, decompose the transition state and leave the enzyme in its unbound sate (Figure 1.9).

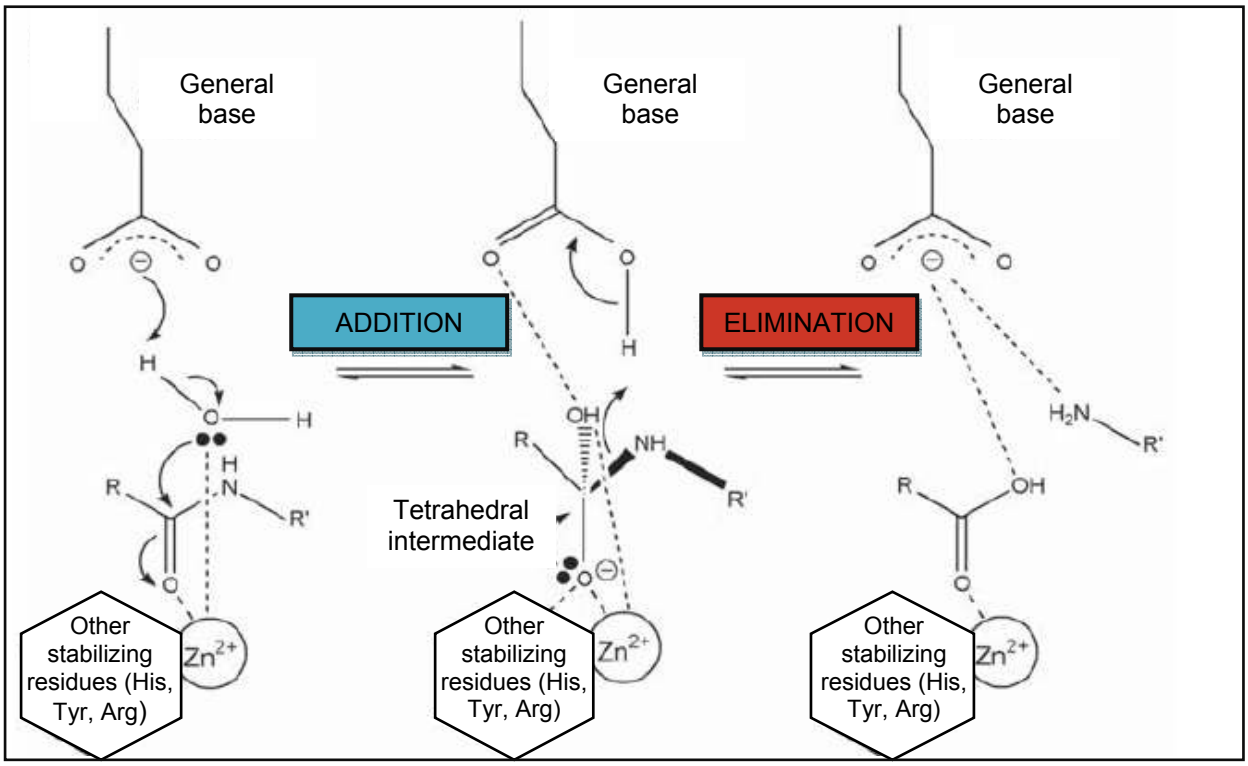

Figure 1.9. The hydrolytic mechanism for metalloproteases, including metzincins (Adapted from Gomis-Rüth, 2003).

1.5.3 Regulation

Since proteases, including the metzincins, catalyse irreversible hydrolytic reactions, strict regulation is a necessity. Protease activity can be controlled *in vivo* by several means, including transcriptional regulation, post-translational modifications and mechanisms which prevent substrate binding to the active site. The latter form of

regulation can occur either through pro-domain inhibition (zymogen activation) and/or blockade of the active site by specific endogenous inhibitors (López-Otín and Bond, 2008). These two regulatory mechanisms form the basis of the following section.

1.5.3.1 Regulation via zymogen activation

Most metzincins are initially synthesised in their latent form (so-called zymogen/ pro-enzyme form) and require proteolytic cleavage to become activated. In the inactive state metzincin activity is blocked by means of a globular pro-domain present in the active site. For the enzyme to be "switched on" the removal of the pro-domain is required. In the pro-domain of leishmanolysins, MMPs and reprolysins the conserved motifs, HRCID, PRCGXPD (Figure 1.10B) and PKMCGV (Figure 1.10C), respectively contain a highly conserved cysteine residue (Figure 1.10) (Barrett *et al.*, 2004). In the latent form, the thiol group of this conserved cysteine substitutes the catalytic water molecule and is coordinated in close proximity with the active site zinc-ion, preventing substrate binding. Dissociation of the cysteine from the zinc-ion "switches" the enzyme from an inactive non-catalytic state to an active catalytic state. This mechanism is known as the "cysteine switch" or "velcro" mechanism and was first characterised in collagenase (Springman *et al.*, 1990).

In vitro, the enzyme can be activated by physical agents such as sodium dodecyl sulfate (SDS) or chaotropic agents, which will unfold the cys-containing pro-peptide and expose the active site. Dissociation can also be established by various reagents that react directly with the thiol group, leaving the cysteine in a non-binding form (Figure 1.10). *In vivo,* activation is mediated by proteolytic enzymes (such as trypsin/ plasmin/ kallikrein), which cleave the pro-peptide, usually ahead of the cysteine. To obtain permanent activated enzyme, the active intermediate enzyme autocatalytically removes the pro-peptide (Figure 1.10) (Woessner, 1991). Astacins, however, do not contain the "cysteine switch" mechanism. These enzymes are not capable of self-activation and require proteolytic cleavage, in order to remove their elongated N-termini (which are blocking substrate binding) from the active site (Bond and Beynon, 1995; Lipscomb and Sträter, 1996). Trypsin-like proteolytic cleavage of the pro-domain constitutes a major mechanism for regulated activation of astacins and will briefly be discussed in Chapter 2.

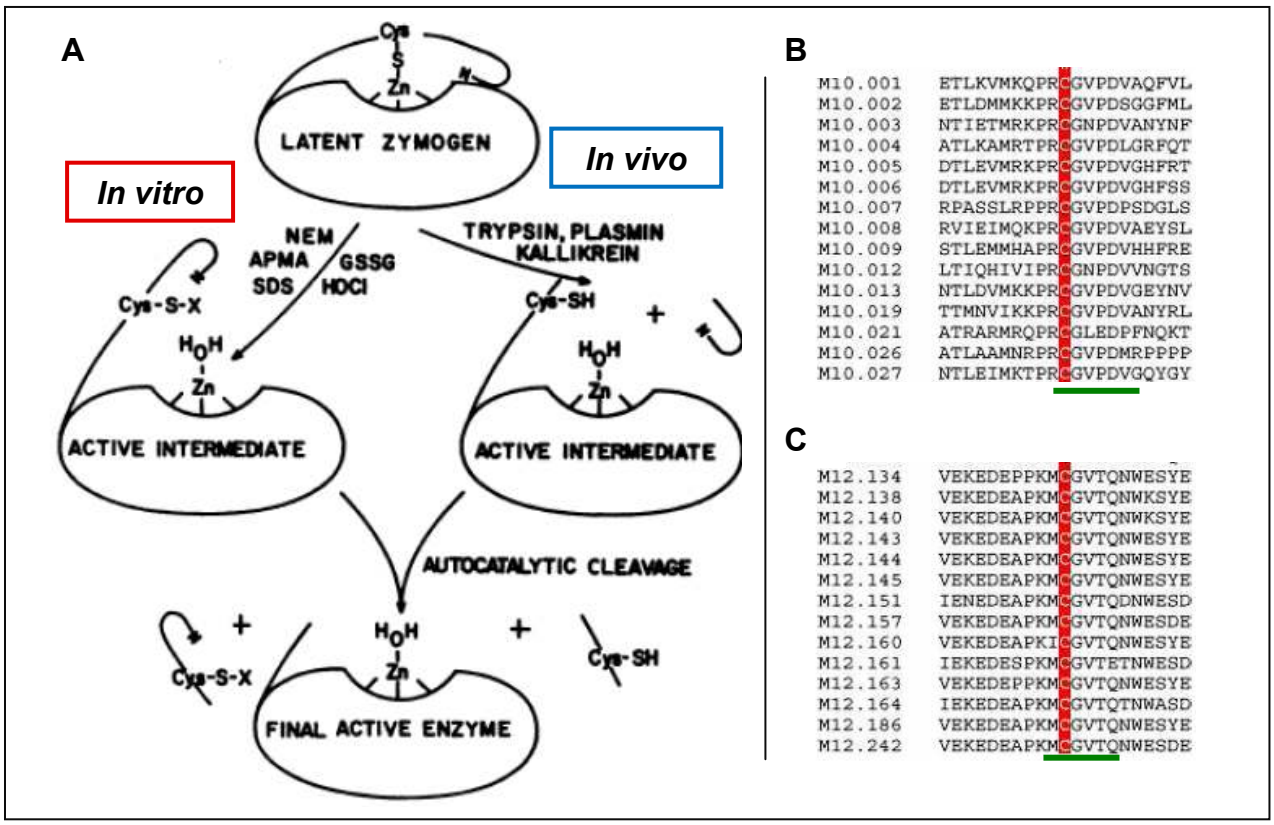

Figure 1.10. The cysteine switch mechanism (A) and alignments showing the conserved cysteine residue in the prodomain of some MMP (B) and reprolysin (C) metzincins (Adapted from Woessner, 1991; Barrett *et al.*, 2004). In figure A both the *in vitro* (left) and *in vivo* (right) activation mechanisms are indicated and the abbreviations correspond to: SDS, sodium dodecyl sulfate; NEM, N-ethylmaleimide; APMA, aminophenylmercuric acetate; GSSG, oxidised glutathione; HOCL, hypochlorous acid. In figure B and C the sequences correspond to various MMP and reprolysins, respectively. The sequences are labeled according to the MEROPS database (http://merops.sanger.ac.uk), for further information see Rawlings *et al.* (2008). The positions of the conserved cysteine and the respective conserved motifs are indicated by the red background and green lines, respectively.

1.5.3.2 Regulation via inhibition

Another very important mechanism for the regulation of metzincin activity involves endogenous inhibitors. The best characterised metzincin metalloprotease inhibitors are the tissue inhibitors of metalloproteases (TIMPs) (Brew and Nagase, 2010). These inhibitors are active against both MMP metzincins and the ADAMs-members of the reprolysins (Edwards *et al.*, 2009), by forming non-covalent binding complexes with the active sites of either latent or activated enzymes. Structural investigations regarding the mode-of-inhibition of stromelysin-1 (MMP3) by TIMP-1, revealed that TIMP chelates the catalytic zinc-ion, expels the water molecule from the active site (Gomis-Rüth *et al.*, 1997) and docks, with its elongated wedge-shape ridge, to the active site cleft in a

substrate manner (Figure 1.8 and 1.11.). *In vitro* metzincins are inhibited by various metal chelators. Due to its zinc-specificity the zinc-chelating compound 1,10-phenanthroline is only required in a low millimolar range, while a broad spectrum chelator, such as ethylenediaminetetraacetic acid (EDTA), is required in high concentrations to be effective. Other metal-chelators include amino acid hydroxymates, dipicoline acid and thiol compounds (Barrett *et al.*, 2004).

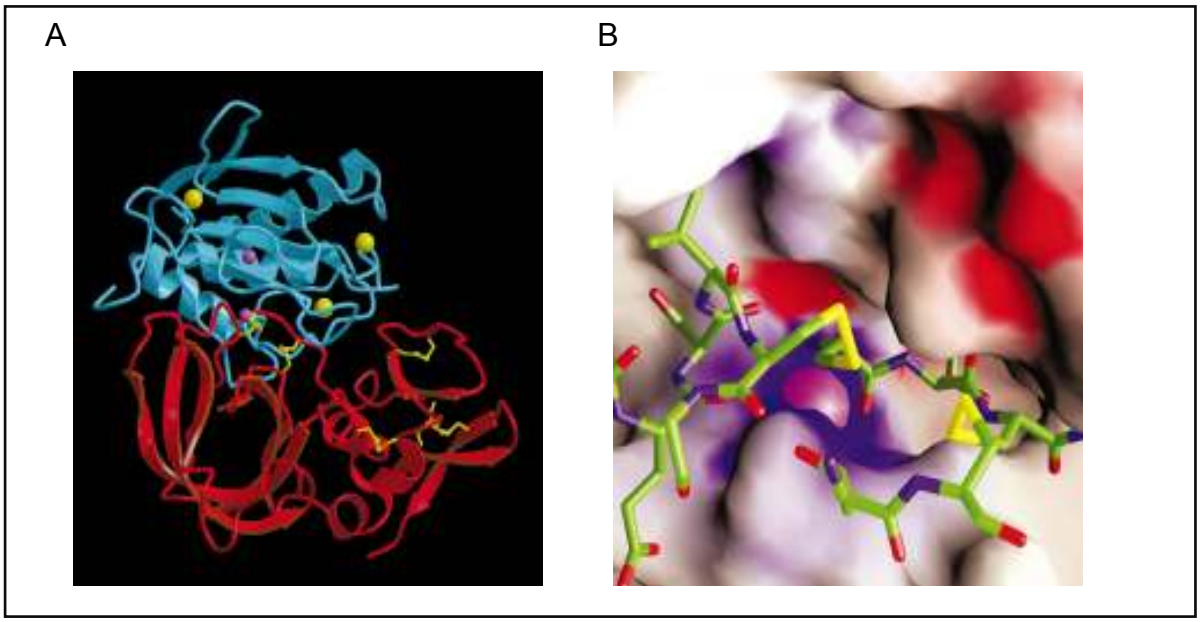

Figure 1.11. A stereo ribbon diagram (A) and a stereo view (B) of the complex formed between TIMP-1 and MMP-3 (Gomis-Rüth *et al.*, 1997). In figure A, the MMP-3 catalytic domain is indicated in blue and TIMP-1 in red; yellow bonds indicate disulphide bridges, and the spheres represent the MMP-3 bound two zinc (pink) and three calcium (yellow) ions. The active-site cleft of MMP-3 is directed halfway towards the bottom. The MMP-3 polypeptide chain starts with a 'cap-like' structure on the left. Cys 1 of TIMP-1 is located on top of the catalytic zinc (centre of the complex). Its first four residues insert between the 'edge' strand (above, front) and the strand blocking off the specificity pocket (below, back) of MMP-3, before the TIMP-1 chain deviates from the active-site cleft to form the N-terminal subdomain (left, bottom) and the C-terminal subdomain (right, bottom). In figure B, the stereo view from the centre of TIMP-1 towards the active-site cleft of MMP-3, which is represented by a solid surface coloured according to the electrostatic potential, is presented. The view is similar to that of figure A. For simplicity, TIMP-1 (with carbons in green, oxygens in red, nitrogens in blue, and sulphurs in yellow) is represented only by the disulphide connected segments Glu 67–Gly 71 (left), Cys 1–Cys 3 (right), and Thr 98–Cys 99 (bottom segment). The N-terminal TIMP-1 residue, Cys 1, is located on top of the catalytic zinc (pink sphere), liganding it through its α-amino and its carbonyl group; Thr 2 extends towards the deep S19 pocket of MMP-3 (right). The TIMP-1 side chains of Ser 68 and Val 69 (left) nestle towards the shallow S2 subsite and the S3 pocket of MMP-3, respectively.

1.6 Metzincins as vaccine targets

Many infectious micro-organisms and parasites require metzincins for survival. Micro-organisms, like vector-based protozoans require these metalloproteases for replication and use them as virulence factors (Gomis-Rüth, 2003; Barrett *et al.*, 2004). Many harmful parasites find metzincin metalloproteases essential for reproduction (Young *et al.*, 2000; Bowles *et al.*, 2008) and host infestation (Francischetti *et al.*, 2003; Gallego *et al.*, 2005; Dzik, 2006). Examples of such metzincins are given in Table 1.3.

Table.1.3. Metzincin metalloproteases which are involved in different vital physiological processes of various protozoans, ecto- and endoparasites.

	Metalloprotease name	Protozoan/ Parasite	Reference
Metzincin virulence factors	Leishmanolysin (GP63)	*Leismania major* (causing cutaneous or dermal leishmaniasis)	(Brittingham *et al.*, 1995)
	Tb-MSP-A Tb-MSP-B Tb-MSP-C	*Trypanosoma brucei* (protozoan parasite causing African sleeping sickness)	(Grandgenett *et al.*, 2007)
Metzincin fusion proteins functioning in host infestation	Ay-MTP-1	*Ancylostoma ceylanicum* (hookworm)	(Williamson *et al.*, 2006)
	Ac-MTP-1	*Ancylostoma caninum* (hookworm)	(Mendez *et al.*, 2005)
Metzincins involved in extracellular coat degradation (egg hatching)	hch-1 MP	*Caenorhabditis elegans* (nematode)	(Hishida *et al.*, 1996)
	Ph-MP(s)	*Pediculus humanus* (body louse)	(Bowles *et al.*, 2008)
	Lc-MP(s)	*Lucilia cuprina* (sheep blowfly)	(Young *et al.*, 2000)
Metzincins involved in parasitic bloodfeeding	IsMP1	*I. scapularis* (ixodid tick)	(Francischetti *et al.*, 2003)
	Metis1 Metis2	*I. ricinus* (ixodid tick)	(Decrem *et al.*, 2007)
	hlESTMP1 hlESTMP2	*H. longicornis* (ixodid tick)	(Harnnoi *et al.*, 2007)

*Abbreviations correspond to: MSP, major surface protease; MTP, astacin-like metalloprotease; MP, metalloprotease; EST, expressed sequence tag.

The vital physiological implications of metzincins in these biological systems make these enzymes attractive targets for vaccines, against these medically and economically important organisms.

Specific metzincins which have been considered as vaccine targets in parasite-borne diseases include: leishmanolysin which is a vital component in the promastigote of the disease causing protozoan, *Leishmania* (Gomis-Rüth, 2003; Barrett *et al.*, 2004); anastacin-like metalloprotease from the hookworm *Ancylostoma caninum,* which is essential for larvae migration and skin penetration (Williamson *et al.*, 2006); and a zinc-metalloprotease from the haemoflagellates parasite, *Cryptobia salmostica*, which acts as a virulence factor (Tan *et al.*, 2008).

1.6.1 Metzincins as anti-tick vaccines

Based on a study of *I. scapularis* tick salivary gland metalloproteases, Francischetti *et al.* proved that reprolysin-like metalloproteases contribute to the development of a fluid blood feeding cavity. The latter is proposed to perform vital anti-haemostatic activities such as fibrinogenolysis and fibrinolysis (Francischetti *et al.*, 2003). Accordingly, Decrem *et al.* proposed that these salivary gland metalloproteases are adequate vaccine candidates, since an elicited immune response would result in a decreased anti-haemostatic ability (Decrem *et al.*, 2007). Subsequent vaccination against these salivary gland metalloproteases did in fact affect the engorgement weight and fecundity of *I. ricinus* (Decrem *et al.*, 2007) as well as *H. longicornis* (Imamura *et al.*, 2009). These reports hold up the notion that tick metzincin metalloproteases are promising anti-tick vaccine candidates and require further investigation to fully explore their potential.

1.7 Aims of this study

This study aims to:

- Identify and invesitigate the expression profiles of reprolysin and astacin-metzincin homologues, from the cattle tick *R. microplus* (Chapter 2).

- Evaluate the phenotype and the differential gene expression profile of the metzincins using *in vivo* RNAi, phenotype assessments and semi-quantitative real-time PCR (Chapter 3).
- Cost effectively express selected domains of the three promising metzincin candidates with the MSP1α-*E. coli* expression system for subsequent cattle vaccine trials (Chapter 4).

Chapter 5 presents a concluding discussion of the scientific contribution of the study, the most important research highlight and the future perspectives.

The knowledge gained from this study has led to the following contributions in scientific journals and at confrence proceedings:

Published, peer-reveiwed international manuscripts:

Manuscripts in progress:

Barnard A.C. and Maritz-Olivier C. (2013). " Cattle vaccine trials using a combinatorial metzincin construct against *Rhipicephalus microplus* infestation. (Manuscript in preparation)

Maritz-Olivier C., Barnard A.C. and van Zyl W.A. (2013). "Selection of reference genes from micro-arrray data for use in standard and quantitative PCR studies in *Rhipichephalus microplus and Rhipichephalus decoloratus*." (Manuscript in preparation)

Published peer-reviewed international manuscripts:

Barnard A.C., Nijhof A.M., Gaspar A.R., Neitz A.W.H., Jongejan F., Maritz-Olivier C. (2012). "Expression profiling, gene silencing and transcriptional networking of metzincin metalloproteases in the cattle tick, *Rhipicephalus (Boophilus) microplus*". Veterinary Parasitology **186**: 403-414.

Barnard A.C., Nijhof A.M., Fick W., Stutzer C. And Maritz-Olivier C (2012). "RNAi in arthropods: Insight into the machinery and applications for understanding the pathogen-vector interface." Genes **3**: 702-741.

Conference proceedings:

Barnard A.C. and Maritz-Olivier C. (2012). "Metzincins: Promising anti-tick vaccines". The second Southern African postgraduate student symposium, Faculty of Veterinary Sciences, Onderstepoort, Pretoria, South Africa.

Barnard A.C. and Maritz-Olivier C (2012). "A reverse genetics approach to evaluate Metzincin metalloproteases as anti-*Rhipichephalus microplus* vaccine candidates" (Poster). South African Genetics and Bioinformatics and Computational Biology Society Conference, Cape Town, South Africa.

Barnard A.C. and Maritz-Olivier C (2012). "A reverse genetics approach to evaluate Metzincin metalloproteases as anti-*Rhipichephalus microplus* vaccine candidates". 1st Regional Conference of the Society for Tropical and Veterinary Medicine (STVM), Phuket, Thailand. Best oral award in the session "Tick and tick-borne diseases".

Barnard A.C. and Maritz-Olivier C. (2012). "A reverse genetics approach to evaluate Metzincin metalloproteases as anti-*Rhipichephalus microplus* vaccine candidates". 23nd South African Society of Biochemistry and Molecular Biology (SASBMB) congress, Drakensberg, Winterton, South Africa

Barnard A.C., Nijhof A.M., Gaspar A.R.M., Neitz A.W.H., Jongejan F., Maritz-Olivier C. (2010). "A reverse genetic approach to evaluate metzincins as anti-*Boophilus microplus* vaccine targets" (Poster). 22nd South African Society of Biochemistry and Molecular Biology (SASBMB) congress, University of the Free State.

1.8 References

Agbede, R. I. S. and D. H. Kemp (1986). "Digestion in the cattle-tick *Boophilus microplus*: Light microscope study of the gut cells in nymphs and females." *International Journal for Parasitology* **15**(2): 147-157.

Almazán, C., R. Lagunes, M. Villar, M. Canales, R. Rosario-Cruz, F. Jongejan and J. de la Fuente (2010). "Identification and characterization of *Rhipicephalus* (*Boophilus*) *microplus* candidate protective antigens for the control of cattle tick infestations." *Parasitology Research* **106**: 471-479.

Aung, K. M., D. Boldbaatar, R. Umemiya-Shirafuji, M. Liao, X. Xuenan, H. Suzuki, R. L. Galay, T. Tanaka and K. Fujisaki (2011). "Scavenger receptor mediates systemic RNA interference in ticks." *PloS one* **6**(12): e28407.

Baffi, M. A., G. R. L. de Souza, C. U. Vieira, C. S. de Sousa, L. R. Gourlart and A. M. Bonetti (2007). "Identification of point mutations in a putative carboxylesterase and their association with acaricide resistance in *Rhipicephalus* (*Boophilus*) *microplus* (Acari: Ixodidae)." *Veterinary Parasitology* **148**(3-4): 301-309.

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

Barrett, A. J., N. D. Rawlings and J. F. Woessner (2004). *Handbook of Proteolytic Enzymes*. Amsterdam, the Netherlands, Elsevier Academic Press.

Becker, A. B. and R. A. Roth (1992). "An unusual active site identified in a family of zinc metalloendopeptidases." *Proceedings of the Natural Academy of Sciences of the United States of America* **89**(9): 3835-3839.

Bellés, X. (2010). "Beyond *Drosophila*: RNAi *in vivo* and functional genomics in insects." *Annual Review of Entomology* **55**: 111 - 128.

Bellgard, M. I., P. M. Moolhuijzen, F. D. Guerrero, D. Schibeci, M. Rodriguez-Valle, D. G. Peterson, S. E. Dowd, R. Barrero, A. Hunter, R. J. Miller and A. E. Lew-Tabor (2012). "CattleTickBase: an integrated Internet-based bioinformatics resource for Rhipicephalus (Boophilus) microplus." *Int J Parasitol* **42**(2): 161-169.

Benjamin, M. A., E. Zhioua and R. S. Ostfeld (2002). "Laboratory and field evaluation of the entomopathogenic fungus *Metarhizium anisopliae* (Deuteromycetes) for controlling questing adult *Ixodes scapularis* (Acari: Ixodidae)." *Journal of Medical Entomology* **39**(5): 723-728.

Berkvens, D. L., D. M. Geysen, G. Chaka, M. Madder and J. R. A. Brandt (1998). "A survey of the ixodid ticks parasitising cattle in the Eastern province of Zambia." *Medical and Veterinary Entomology* **12**(3): 234-240.

Beugnet, F. and L. Chardonnet (1995). "Tick resistance to pyrethroids in New Caledonia." *Veterinary Parasitology* **56**(4): 325-338.

Blobel, C. P. (2005). "ADAMs: Key components in egfr signalling and development." *Nature Reviews Molecular Cell Biology* **6**(1): 32-43.

Bock, R. E., L. A. Jackson, A. de Vos and W. K. Jorgensen (2004). "Babesiosis of cattle." *Parasitology* **129**: S247-S269.

Bode, W., F.-X. Gomis-Rüth, R. Huber, R. Zwilling and W. Stöcker (1992). "Structure of astacin and implications for activation of astacins and zinc-ligation of collagenases." *Nature* **358**(6382): 164-167.

Bode, W., F.-X. Gomis-Rüth and W. Stöckler (1993). "Astacins, serralysins, snake venom and matrix mettalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn)

and topologies and should be grouped into a common family, the 'metzincins'." *Federation of European Biochemical Societies Letters* **331**(1-2): 134-140.

Bode, W., P. Reinemer, R. Huber, T. Kleine, S. Schmierer and H. Tschesche (1994). "The X-ray crystal structure of the catalytic domain of human neutrophil collagenase inhibited by a substrate analogue reveals the essentials for catalysis and specificity." *EMBO Journal* **13**(6): 1263-1269.

Bond, J. S. and R. J. Beynon (1995). "The astacin family of metalloendopeptidases." *Protein Science* **4**(7): 1247-1261.

Bowles, V. M., A. R. Young and S. C. Barker (2008). "Mettalloproteases and egg-hatching in *Pediculus humanus*, the body (clothes) louse of humans (Phthiraptera: Insecta)." *Parasitology* **135**(1): 125-130.

Brew, K. and H. Nagase (2010). "The tissue inhibitors of metalloproteinases (TIMPs): An ancient family with structural and functional diversity." *Biochim Biophys-Acta* **1803**: 55–71.

Brittingham, A., C. J. Morrison, W. R. McMaster, B. S. McGwire, K.-P. Chang and D. M. Mosser (1995). "Role of the *Leishmania* surface protease gp63 in complement fixation, cell adhesion, and resistance to complement-mediated lysis." *Journal of Immunology* **155**(6): 3102-3111.

Calixto, A., D. Chelur, I. Topalidou, X. Chen and M. Chalfie (2010). "Enhanced neuronal RNAi in *C. elegans* using SID-1." *Nature Methods* **7**(7): 554 - 559.

Carthew, R. W. and E. J. Sontheimer (2009). "Origins and mechanisms of miRNAs and siRNAs." *Cell* **136**: 642 - 655.

Chang, C. and Z. Werb (2001). "The many faces of metalloproteases: Cell growth, invasion, angiogenesis and metastasis." *Trends in Cell Biology* **11**(11): S37-S43.

Chen, A. C., H. He and R. Davey (2007). "Mutations in a putative octopamine receptor gene in amitraz-resistant cattle ticks." *Veterinary Parasitology* **148**: 379-383.

Cheng, Y., H. Wu and D. Li (1999). "An inhibitor selective for collagen-stimulated platelet aggregation from the salivary glands of the hard tick *Haemaphysalis longicornis* and its mechanism of action." *Science in China Series C Life Sciences* **42**(5): 463-464.

Chernov, A. V., S. Baranovskaya, V. S. Golubkov, D. R. Wakeman, E. Y. Snyder, R. Williams and A. Y. M.-b. t. a. e. p. o. m. m. Strongin, collagens and related genes in cancer. (2010). "Microarray-based transcriptional and epigenetic profiling of matrix metalloproteinases, collagens and related genes in cancer." *Journal of Biological Chemistry* **285**: 19647-19659.

Das, S., G. Banerjee, K. DePonte, N. Marcantonio, F. S. Kantor and E. Fikrig (2001). "Salp25D, an *Ixodes scapularis* antioxidant, is 1 of 14 immunodominant antigens in engorged tick salivary glands." *Journal of Infectious Diseases* **184**(8): 1056-1064.

Davie, E. W. and H. Neurath (1955). "Identification of a peptide released during autocatalytic activation of trypsinogen." *Journal of Biological Chemistry* **212**(2): 515-529.

de la Fuente, J., C. Almazán, M. Canales, J. M. P. de la Lastra, K. M. Kocan and P. Willadsen (2007a). "A ten-year review of commercial vaccine performance for control of tick infestations on cattle." *Animal Health Research Reviews* **8**(1): 23-28.

de la Fuente, J. and K. M. Kocan (2006). "Strategies for the development of vaccines for control of ixodid tick species." *Parasite Immunology* **28**: 275-283.

de la Fuente, J., M. Rodríguez, M. Redondo, C. Montero, J. C. Garcia-Garcia, L. Méndez, E. Serrano, M. Valdés, A. Enriquez, M. Canales, E. Ramos, O. Boué, H. Machado, R. Lleonart, C. A. de Armas, S. Rey, J. L. Rodríguez, M. Artiles and L. Garcia (1998). "Field studies and cost-effectiveness analysis of vaccination with Gavac™ against the cattle tick *Boophilus microplus*." *Vaccine* **16**(4): 366-373.

de Rose, R., R. V. McKenna, G. Cobon, J. Tennent, H. Zakrzewski, K. Gale, P. R. Wood, J. P. Scheerlinck and P. Willadsen (1999). "Bm86 antigen induces a protective immune response against *Boophilus microplus* following DNA and protein vaccination in sheep." *Veterinary Immunology and Immunopathology* **71**(3-4): 151-160.

Decrem, Y., M. Mariller, K. Lahaye, V. Blasioli, J. Beaufays, K. Z. Boudjeltia, M. Vanhaeverbeek, M. Cérutti, L. Vanhamme and E. Godfroid (2007). "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**(5): 549-560.

Dickinson, R. G., J. E. O'Hagan, M. Schotz, K. C. Binnington and M. P. Hegarty (1976). "Prostaglandin in the saliva of the cattle tick *Boophilus microplus*." *Australian Journal of Experimental Biological and Medical Sciences* **54**(5): 475-486.

Dzik, J. M. (2006). "Molecules released by helminth parasites involved in host colonization." *Acta Biochimica Polonica* **53**(1): 33-64.

Edwards, D. R., M. M. Handsley and C. J. Pennington (2009). "The ADAM metalloproteinases." *Molecular Aspects of Medicine* **29**(5): 258-289.

El Shoura, S. M. (1988). "Ultrastructural studies on the midgut epithelium and digestion in the female tick *Argas* (*Persicargas*) *arboreus* (Ixodoidea: Argasidae)." *Experimental and Applied Acarology* **5**(1-2): 121-136.

Elvin, C. M. and D. H. Kemp (1994). "Generic approaches to obtaining efficacious antigens from vector arthropods." *International Journal for Parasitology* **24**(1): 67–79.

Endres, K. and F. Fahrenholz (2010). "Upregulation of the α-secretase ADAM10 – risk or reason for hope?" *FEBS J* **277**: 1585–1596

Estrada-Peña, A., A. Bouattour, J.-L. Camicas, A. Guglielmone, I. Horak, F. Jongejan, A. Latif, R. Pegram and A. R. Walker (2006). "The known distribution and ecological preferences of the tick subgenus *Boophilus* (Acari: Ixodidae) in Africa and Latin America." *Experimental and Applied Acarology* **38**(2-3): 219-235.

Francischetti, I. M. B., T. N. Mather and J. M. C. Ribeiro (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**(4): 869-875.

Francischetti, I. M. B., T. N. Mather and J. M. C. Ribeiro (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**(5): 886-898.

Francischetti, I. M. B., J. G. Valenzuela, J. F. Andersen, T. N. Mather and J. M. C. Ribeiro (2002). "Ixolaris, a novel recombinant tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick, *Ixodes scapularis*: Identification of factor X and factor Xa as scaffolds for the inhibition of factor VIIa/tissue factor complex." *Blood* **99**(10): 3602-3612.

Frazzon, A. P. G., I. D. S. Vaz, Jr., A. Masuda, A. Schrank and M. H. Valnstein (2000). "*In vitro* assessment of *Metarhizium anisopliae* isolates to control the cattle tick *Boophilus microplus*." *Veterinary Parasitology* **94**(1-2): 117-125.

Friedhoff, K. T. and R. D. Smith (1981). *Transmission of Babesia by ticks*. New York, USA, Academic Press.

Gallego, S. G., A. Loukas, R. W. Slade, F. A. Neva, R. Varatharajalu, T. B. Nutman and P. J. Brindley (2005). "Identification of an astacin-like metallo-proteinase transcript from the infective larvae of *Strongyloides stercoralis*." *Parasitology International* **54**(2): 123-133.

García-García, J. C., I. L. Gonzalez, D. M. González, M. Valdés, L. Méndez, J. Lamberti, B. D'Agostino, D. Citroni, H. Fragoso, M. Ortiz, M. Rodríguez and J. de la Fuente (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.

Gindin, G., M. Samish, G. Zangi, A. Mishoutchenko and I. Glazer (2002). "The susceptibility of different species and stages of ticks to entomopathogenic fungi." *Experimental and Applied Acarology* **28**(1-4): 283-288.

Gomis-Rüth, F.-X. (2003). "Structural aspects of the metzincin clan of metalloendopeptidases." *Molecular Biotechnology* **24**(2): 157-202.

Gomis-Rüth, F.-X. (2009). "Catalytic domain architecture of metzincin metalloproteases." *Journal of Biological Chemistry* **284**(23): 15353-15357.

Gomis-Rüth, F.-X., K. Maskos, M. Betz, A. Bergner, R. Huber, K. Suzuki, N. Yoshida, H. Nagase, K. Brew, G. P. Bourenkov, H. Bartunik and W. Bode (1997). "Mechanism of inhibition of the human matrix metalloproteinase stromelysin-1 by TIMP-1." *Nature* **389**(6646): 77-81.

Gomis-Rüth, F.-X., W. Stöcker, R. Huber, R. Zwilling and W. Bode (1993b). "Refined 1.8 Å X-ray crystal structure of astacin, a zinc-endopeptidase from the crayfish *Astacus astacus* L.: Structure determination, refinement, molecular structure and comparison with thermolysin." *Journal of Molecular Biology* **229**(4): 945-968.

Graf, J.-F., R. Gogolewski, N. Leach-Bing, G. A. Sabatini, M. B. Molento, E. L. Bordin and G. J. Arantes (2004). "Tick control: An industry point of view." *Parasitology* **129**: S427-S442.

Grams, F., P. Reinemer, J. C. Powers, T. Kleine, M. Pieper, H. Tschesche, R. Huber and W. Bode (1995). "X-ray structures of human neutrophil collagenase complexed with peptide hydroxamate and peptide thiol inhibitors. Implications for substrate binding and rational drug design." *European Journal of Biochemistry* **228**(3): 830-841.

Grandgenett, P. M., K. Otsu, H. R. Wilson, M. E. Wilson and J. E. Donelson (2007). "A function for a specific zinc metalloprotease of African trypanosomes." *PLoS Pathogens* **3**(10): 1432-1445.

Guerrero, F. D., L. Lovis and J. R. Martins (2012a). "Acaricide resistance mechanisms in *Rhipicephalus (Boophilus) microplus*." *Rev. Bras. Parasitol. Vet.* **21**: 1-6.

Guglielmone, A. A., R. G. Robbins, D. A. Apanaskevich, T. N. Petney, A. Estrada-Peña, I. G. Horak, R. Shao and S. C. Barker (2010). "The Argasidae, Ixodidae and Nuttalliellidae (Acari: Ixodida) of the world: A list of valid species names." *Zootaxa* **2528**.

Harnnoi, T., T. Sakaguchi, Y. Nishikawa, X. Xuan and K. Fujisaki (2007). "Molecular characterization and comparative study of 6 salivary gland metalloproteases from the tick, *Haemaphysalis longicornis*." *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology* **147**(1): 93-101.

Hernandez, R., H. He, A. C. Chen, S. D. Waghela, G. W. Ivie, J. E. George and G. G. Wagner (2000). "Identificiation of a point mutation in an esterase gene in different populations of the southern cattle tick, *Boophilus microplus*." *Insect Biochemistry and Molecular Biology* **30**(10): 969-977.

Hishida, R., T. Ishihara, K. Kondo and I. Katsura (1996). "hch-1, a gene required for normal hatching and normal migration of a neurblast in *C. elegans*, encodes a protein related to TOLLOID and BMP-1." *EMBO Journal* **15**(16): 4111-4122.

Hoogstraal, H. (1956). *African Ixodoidea. I. Ticks of the Sudan.* Washington DC, USA, United States Navy Bureau of Medicine and Surgery.

Hooper, N. M. (1994). "Families of zinc metalloproteases." *Federation of European Biochemical Societies Letters* **354**(1): 1-6.

Hope, M., M. Menzies and D. Kemp (2010). "Identification of a Dieldrin Resistance-Associated Mutation in *Rhipicephalus (Boophilus) microplus* (Acari: Ixodidae)." *Journal of Economic Entomology* **103**: 1355-1359.

Horn, M., M. Nussbaumerova, M. Sanda, Z. Kovarova, J. Srba, Z. Franta, D. Sojka, M. Bogyo, C. R. Caffrey, P. Kopacek and M. Mares (2009). "Hemoglobin digestion in blood-feeding ticks: mapping a multipeptidase pathway by functional proteomics." *Chemistry & biology* **16**(10): 1053-1063.

Imamura, S., I. da Silva Vaz, Jr., S. Konnai, S. Yamada, C. Nakajima, M. Onuma and K. Ohashi (2009). "Effect of vaccination with a recombinant metalloprotease from *Haemaphysalis longicornis*." *Experimental and Applied Acarology* **48**(4): 345-358.

Jamroz, R. C., F. D. Guerrero, J. H. Pruett, D. D. Oehler and R. J. Miller (2000). "Molecular and biochemical survey of acaricide resistance mechanisms in larvae from Mexican strains of the southern cattle tick, *Boophilus microplus*." *Journal of Insect Physiology* **46**(5): 685-695.

Jaworski, D. C., A. Jasinskas, C. N. Metz, R. Bucala and A. G. Barbour (2001). "Identification and characterization of a homologue of the pro-inflammatory cytokine Macrophage Migration Inhibitory Factor in the tick, *Amblyomma americanum*." *Insect Molecular Biology* **10**(4): 323–331.

Johnston, L. A. Y., D. D. Kemp and R. D. Pearson (1986). "Immunization of cattle against *Boophilus microplus* using derived from adult female ticks: Effects of induced immunity on tick population. ." *International Journal of Parasitology* **16**: 27-34.

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

Jongeneel, C. V., J. Bouvier and A. Bairoch (1989). "A unique signature identifies a family of zinc-dependent metallopeptidases." *Federation of European Biochemical Societies Letters* **242**(2): 211-214.

Jonsson, N. N. (1997). "Control of cattle ticks (*Boophilus microplus*) on Queensland dairy farms." *Australian Veterinary Journal* **75**(11): 802-807.

Jose, A. M., J. J. Smith and C. P. Hunter (2009). "Export of RNA silencing from *C. elegans* tissues does not require the RNA channel SID-1." *Proceedings of the National Academy of Sciences of the United States of America* **106**(7): 2283 - 2288.

Katsande, T. C., W. Mazhowu, J. A. Turton and D. Munodzana (1996). "*Babesia bovis* case reports and the current distribution of *Boophilus microplus* in Zimbabwe." *Zimbabwe Veterinary Journal* **27**(1): 33-36.

Kjemtrup, A. M. and P. A. Conrad (2000). "Human babesiosis: An emerging tick-borne disease." *International Journal for Parasitology* **30**(12-13): 1323-1337.

Klafke, G. M., E. Castro-Janer, M. C. Mendes, A. Namindome and T. T. S. Schumaker (2012). "Applicability of in vitro bioassays for the diagnosis of ivermectin resistance in Rhipicephalus microplus (Acari: Ixodidae)." *Veterinary Parasitology* **184**: 212-220.

Kunz, S. E. and D. H. Kemp (1994). "Insecticides and acaricides: Resistance and environmental impact." *OIE Revue Scientifique et Technique* **13**(4): 1249-1286.

Lai, R., H. Takeuchi, J. Jonczy, H. H. Rees and P. C. Turner (2004). "A thrombin inhibitor from the ixodid tick, *Amblyomma hebraeum*." *Gene* **342**(2): 243-249.

Li, A. Y., A. C. Chen, R. J. Miller, R. B. Davey and J. E. George (2007). "Acaricide resistance and synergism between permethrin and amitraz against susceptible and resistant strains of *Boophilus microplus* (Acari: Ixodidae)." *Pest Management Science* **63**(9): 882-889.

Li, A. Y., R. B. Davey, R. J. Miller and J. E. George (2003). "Resistance to coumaphos and diazinon in *Boophilus microplus* (Acari: Ixodidae) and evidence for the involvement of an oxidative detoxification mechanism." *Journal of Medical Entomology* **40**(4): 482-490.

Li, A. Y., R. B. Davey, R. J. Miller and J. E. George (2004). "Detection and characterization of amitraz resistance in the southern cattle tick, *Boophilus microplus* (Acari:Ixodidae)." *Journal of Medical Entomology* **41**(2): 193-200.

Lipscomb, W. N. and N. Sträter (1996). "Recent advances in zinc enzymology." *Chemical Reviews* **96**(7): 2375-2433.

López-Otín, C. and J. S. Bond (2008). "Proteases: Multifunctional enzymes in life and disease." *Journal of Biological Chemistry* **283**(45): 30433-30437.

Lynen, G., P. Zeman, C. Bakuname, G. Di Giulio, P. Mtui, P. Sanka and F. Jongejan (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**(2): 147-164.

Macedo-Ribeiro, S., C. Almeida, B. M. Calisto, T. Friedrich, R. Mentele, J. Stürzebecher, P. Fuentes-Prior and P. J. Pereira (2008). "Isolation, cloning and structural characterisation of boophilin, a multifunctional Kunitz-type proteinase inhibitor from the cattle tick." *PloS one* **3**(2): e1624.

Macleod, J. and B. Mwanaumo (1978). "Ecological studies on ixodid ticks (Acari: Ixodidae) in Zambia. IV. Some anomalous infestation patterns in the northern and eastern regions." *Bulletin of Entomological Research* **68**(3): 409-429.

Madder, M., E. Thys, L. Achi, A. Toure and R. De Deken (2011). "Rhipicephalus (Boophilus) microplus: a most successful invasive tick species in West-Africa." *Experimental & applied acarology* **53**(2): 139-145.

Madder, M., E. Thys, D. Geysen, C. Baudoux and I. Horak (2007). "*Boophilus microplus* ticks found in West Africa." *Experimental and Applied Acarology* **43**(3): 233-234.

Maritz-Olivier, C., C. Stutzer, F. Jongejan, A. W. H. Neitz and A. R. M. Gaspar (2007). "Tick anti-hemostatics: Targets for future vaccines and therapeutics." *Trends in Parasitology* **23**(9): 397-407.

Mason, C. A. and R. A. I. Norval (1980). "The ticks of Zimbabwe. I. The genus *Boophilus*." *Zimbabwe Veterinary Journal* **11**(3): 36-43.

McQuibban, G. A., J.-H. Gong, E. M. Tam, C. A. G. McCulloch, I. Clark-Lewis and C. M. Overall (2000). "Inflammation dampened by gelatinase A cleavage of monocyte chemoattractant protein-3." *Science* **289**(5484): 1202-1206.

Mendez, S., B. Zhan, G. Goud, K. Ghosh, A. Dobardzic, W. Wu, S. Liu, V. Deumic, R. Dobardzic, Y. Liu, J. Bethony and P. J. Hotez (2005). "Effect of combining the larval antigens *Ancylostoma* secreted protein 2 (ASP-2) and metalloprotease 1 (MTP-1) in protecting hamsters against hookworm infection and disease caused by *Ancylostoma ceylanicum*." *Vaccine* **23**(24): 3123-3130.

Miller, R. J., R. B. Davey and J. E. George (1999). "Characterization of pyrethroid resistance and susceptibility to coumaphos in Mexican *Boophilus microplus* (Acari: Ixodidae)." *Journal of Medical Entomology* **36**(5): 533-538.

Moolhuijzen, P. M., A. E. Lew-Tabor, J. A. Morgan, M. R. Valle, D. G. Peterson, S. E. Dowd, F. D. Guerrero, M. I. Bellgard and R. Appels (2011). "The complexity of Rhipicephalus (Boophilus) microplus genome characterised through detailed analysis of two BAC clones." *BMC research notes* **4**: 254.

Morgan, J. A., S. W. Corley, L. A. Jackson, A. E. Lew-Tabor, P. M. Moolhuijzen and N. N. Jonsson (2009). "Identification of a mutation in the para-sodium channel gene of the cattle tick Rhipicephalus (Boophilus) microplus associated with resistance to synthetic pyrethroid acaricides." *Int J Parasitol* **39**(7): 775-779.

Motoyashiki, T., A. T. Tu, D. A. Azimov and K. Ibragim (2003). "Isolation of anticoagulant from the venom of tick, *Boophilus calcaratus*, from Uzbekistan " *Thrombosis Research* **110**(4): 235-241.

Mulenga, A., K. R. Macaluso, J. A. Simser and A. F. Azad (2003). "The American dog tick, *Dermacentor variabilis*, encodes a functional histamine release factor homolog." *Insect Biochemistry and Molecular Biology* **33**(9): 911-919.

Narasimhan, S., R. A. Koski, B. Beaulieu, J. F. Anderson, N. Ramamoorthi, F. Kantor, M. Cappello and E. Fikrig (2002). " A novel family of anticoagulants from the saliva of *Ixodes scapularis*." *Insect Molecular Biology* **11**(6): 641-650.

Neurath, H. and K. A. Walsh (1976). "Role of proteolytic enzymes in biological regulation (A review)." *Proceedings of the Natural Academy of Sciences of the United States of America* **73**(11): 3825-3832.

Paesen, G. C., P. L. Adams, K. Harlos, P. A. Nuttall and D. I. Stuart (1999). "Tick histamine-binding proteins: Isolation, cloning and three-dimensional structure." *Molecular Cell* **3**(5): 661-671.

Peter, R. J., P. van den Bossche, B. L. Penzhorn and B. Sharp (2005). "Tick, fly, and mosquito control - Lessons from the past, solutions for the future." *Veterinary Parasitology* **132** (3-4): 205-215.

Pound, J. M., J. E. George, D. M. Kammlah, K. H. Lohmeyer and R. B. Davey (2010). "Evidence for the role of white-tailed deer (Artiodactyla: Cervidae) in epizootiology of cattle ticks and southern cattle ticks (Acari: Ixodidae) in reinfestations along the Texas/Mexico border in South Texas: A review and update." *Journal of Economic Entomology* **103**(2): 211 - 218.

Rawlings, N. D., F. R. Morton, C. Y. Kok, J. Kong and A. J. Barrett (2008). "MEROPS: The peptidase database." *Nucleic Acids Research* **36**: D320-D325.

Ribeiro, J. M. C., P. M. Evans, J. L. MacSwain and J. Sauer (1992). "*Amblyomma americanum*: Characterization of salivary prostaglandins E2 and F2 α by RP-HPLC/bioassay and gas chromatography-mass spectrometry. ." *Experimental Parasitology* **74**(1): 112-116.

Ribeiro, J. M. C., G. T. Makoul, J. Levine, D. R. Robinson and A. Spielman (1985). "Antihemostatic, antiinflammatory, and immunosuppressive properties of the saliva of a tick, *Ixodes dammini*." *Journal of Experimental Medicine* **161**(2): 332-344.

Rodríguez, M. M., J. Bisset, M. Ruiz and A. Soca (2002). "Cross-resistance to pyrethroid and organophosphorus insecticides induced by selection with temephos in *Aedes aegypti* (Diptera: Culicidae) from Cuba." *Journal of Medical Entomology* **39**(6): 882-888.

Rosario-Cruz, R., F. D. Guerrero, R. J. Miller, R. I. Rodriguez-Vivas, M. Tijerina, D. I. Domínguez-Garcia, R. Hernandez-Ortiz, A. J. Cornel, R. D. McAbee and M. A. Alonso-Diaz (2009). "Molecular survey of pyrethroid resistance mechanisms in Mexican field populations of *Rhipicephalus (Boophilus) microplus*. ." *Parasitology Research* **105**: 1145-1153.

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

Springman, E. B., E. L. Angleton, H. Birkedal-Hansen and H. E. van Wart (1990). "Multiple modes of activation of latent human fibroblast collagenase: Evidence for the role of a Cys73 active-site zinc complex in latency and a 'cysteine switch' mechanism for activation." *Proceedings of the Natural Academy of Sciences of the United States of America* **87**(1): 364-368.

Sterchi, E. E., W. Stöcker and J. S. Bond (2009). "Meprins, membrane-bound and secreted astacin metalloproteinases." *Molecular Aspects of Medicine* **29**(5): 309-328.

Stöcker, W. and W. Bode (1995). "Structural features of a superfamily of zinc-endopeptidases: The metzincins." *Current Opinion in Structural Biology* **5**(3): 383-390.

Stöcker, W., F. Grams, U. Baumann, P. Reinemer, F.-X. Gomis-Rüth, D. B. McKay and W. Bode (1995). "The metzincins - Topological and sequential relations between the astacins, adamalysins, serralysins and matrixins (collagenases) define a superfamily of zinc-peptidases." *Protein Science* **4**(5): 823-840.

Tan, C.-W., P. Jesudhasan and P. T. K. Woo (2008). "Towards a metalloprotease-DNA vaccine against piscine cryptobiosis caused by *Cryptobia salmostica*." *Parasitology Research* **102**(2): 265-275.

Tønnesen, M. H., B. L. Penzhorn, N. R. Bryson, W. H. Stoltsz and T. Masibigiri (2004). "Displacement of *Boophilus decoloratus* by *Boophilus microplus* in the Soutpansberg region, Limpopo Province, South Africa." *Experimental and Applied Acarology* **32**(3): 199-208.

Trager, W. (1939). "Acquired immunity to ticks." *The Journal of Parasitology* **25**: 57-81.

Turk, B. (2006). "Targeting proteases: Successes, failures and future prospects." *Nature Reviews Drug Discovery* **5**(9): 785-799.

Uilenberg, G. (2006). "Babesia - A historical overview." *Veterinary Parasitology* **138**(1-2): 3-10.

Vallee, B. L. and D. S. Auld (1990). "Active-site ligands and activated H_2O of zinc enzymes." *Proceedings of the Natural Academy of Sciences of the United States of America* **87**(1): 220-224.

Villarino, M. A., S. D. Waghela and G. G. Wagner (2003). "Biochemical detection of esterases in the adult female integument of organophosphate-resistant *Boophilus microplus* (Acari: Ixodidae)." *Journal of Medical Entomology* **40**(1): 52-57.

Wang, H. and P. A. Nuttall (1999). "Immunoglobulin-binding proteins in ticks: New target for vaccine development against a blood-feeding parasite." *Cellular and Molecular Life Sciences* **56**(3-4): 286-295.

Wang, X., L. B. Coons, D. B. Taylor, S. E. Stevens, Jr. and T. K. Gartner (1996). "Variabilin, a novel RGD-containing antagonist of glycoprotein IIb-IIIa and platelet aggregation inhibitor from the hard tick *Dermacentor variabilis*." *Journal of Biological Chemistry* **271**(30): 17785-17790.

Wedderburn, P. A., T. D. Jagger, B. McCartan and A. G. Hunter (1991). "Distribution of *Boophilus* species ticks in Swaziland." *Tropical Animal Health and Production* **23**(3): 167-171.

White, N., R. W. Sutherst, N. Hall and P. Whish-Wilson (2003). "The vulnerability of the Australian beef industry to impacts of the cattle tick (*Boophilus microplus*) under climate change." *Climatic Change* **61**(1-2): 157-190.

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

Willadsen, P. (2006). "Tick control: Thoughts on a research agenda." *Veterinary Parasitology* **138**(1-2): 161-168.

Williamson, A. L., S. Lustigman, Y. Oksov, V. Deumic, J. Plieskatt, S. Mendez, B. Zhan, M. E. Bottazzi, P. J. Hotez and A. Loukas (2006). "*Ancylostoma caninum* MTP-1, an astacin-like mettaloprotease secreted by infective hookworm larvae, is involved in tissue migration." *Infection and Immunity* **74**(2): 961-967.

Woessner, J. F., Jr. (1991). "Matrix metalloproteinases and their inhibitors in connective tissue remodeling." *FASEB Journal* **5**(8): 2145-2154.

Yiallouros, I., S. Vassiliou, A. Yiotakis, R. Zwilling, W. Stöcker and V. Dive (1998). "Phosphinic peptides, the first potent inhibitors of astacin, behave as extremely slow-binding inhibitors." *Biochemical Journal* **331**(2): 375-379.

Yoon, K. S., J.-R. Gao, S. H. Lee, G. C. Coles, T. L. Meinking, D. Taplin, J. D. Edman, M. Takano-Lee and M. Clark (2004). "Resistance and cross-resistance to insecticides in human head lice from Florida and California." *Pesticide Biochemistry and Physiology* **80**(3): 192-201.

Young, A. R., N. Mancuso, E. N. T. Meeusen and V. M. Bowles (2000). "Characterisation of proteases involved in egg hatching of the sheep blowfly, *Lucilia cuprina*." *International Journal for Parasitology* **30**(8): 925-932.

Chapter 2

Identification, cloning and expression profiling of astacin and reprolysin homologues from the hard tick, *R. microplus*

2.1 Introduction

2.1.1 The astacin family (MEROPS classification: clan MA(M), family M12A)

This family was named after its first member, astacin (EC 3.4.24.21) (Figure 2.1), a digestive enzyme of the crayfish *Astacus astacus L.* (Bond and Beynon, 1995). To date, more than 200 members of this family have been identified in organisms ranging from bacteria to mammalians, but none have been found in plants or fungi (Gomis-Rüth, 2009). This family comprises of a diverse range of members including: Meprins, bone morphogenetic protein-1 (BMP-1), *tolloids* and mammal *tolloid*-like proteins, *Hydra vulgaris* metalloprotease (HMP), sea urchin blastula protein-10 (BP 10), nephrosin, *Strongylocentrotus purpuratus* astacin (SPAN) and the “hatching” enzymes alveolin, ovostacin, ‘low’ and ‘high’ chariolytic enzymes (LCE and HCE) and QuCAM-1. Like prototype astacin, most astacin-family members contain a prepro segment which directs the proteins into the endoplasmic reticulum (ER) during biosynthesis (Bond and Beynon, 1995). This is consistent with the finding that most proteins of this family are secreted or plasma membrane bound (Sterchi *et al.*, 2009).

2.1.1.1 Distinguishing features

The structure of astacin from the crayfish *A. astacus L.*, which was the first metzincin structure to be solved, together with topological comparative studies on other astacin-like proteins, revealed several distinguishing characteristics of this family (Bode *et al.*, 1992). The overall structure of astacin exhibits a Pacman-like spherical shape: it is composed of two subdomains, separated by the active site cleft in the middle (Figure 2.1). Since most astacins comprise protruding lower domains, their active sites have long, deep incised substrate anchoring grooves, which accommodate seven or more amino acids. In the active site cleft of astacins, the catalytic zinc-ion is triogonal-

bipyramidal coordinated, due to the additional fifth zinc-ligand: a tyrosine residue present in the conserved “Met-turn” motif, SBMHY (Figure 2.1). Since the liganding Oη atom of the tyrosine-residue is slightly further from the zinc-ion (2.5 Å), it is still not absolutely clear if it serves as a zinc-binding ligand. However, the phenolic side chain is within hydrogen bonding distance and is able to flip back and forth in a motion referred to as “tyrosine switching” during substrate binding. This allows stabilisation of the enzyme-substrate complex and/or transition state intermediate (Park and Ming, 1998; Gomis-Rüth, 2003). Other remarkable common structural features include: a hydrophobic cluster in the N-terminal subdomain, a conserved solvent-filled cavity buried in the C-terminal and two intramolecular disulphide bridges, which are formed by four conserved cysteine residues.

While mature *A. astacus L.* astacin contains only the catalytic metzincin domain (HEBXHA/VBGFXH), most other family members are multidomain proteins. In addition to the catalytic domain, they have extra C-terminal domains. Based on the C-terminal domain composition, astacins have been classified into 3 major subgroups: the Meprin proteases, *Tolloid*/BMP-1 proteases, and the “hatching enzymes”.

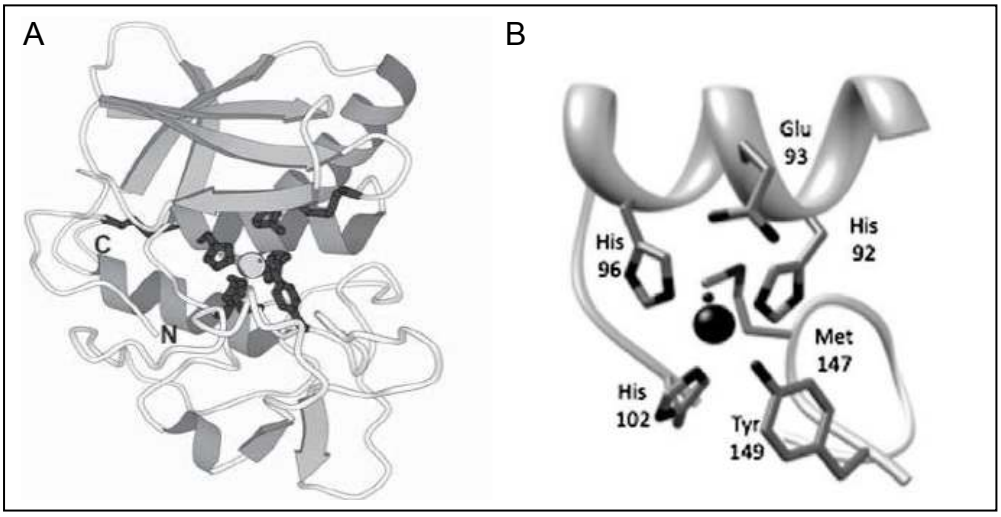

Figure 2.1. Richardson diagram of astacin (PDB 1ast) (A) and a schematic zoomed-in view of the active site (B) (Gomis-Rüth, 2003; Sterchi *et al.*, 2009). In figure A: prototype astacin is shown in the standard orientation (Bode *et al.*, 1992; Gomis-Rüth *et al.*, 1993a), that is, viewing into the active-site groove, which runs horizontally from left to right. The N- and C-termini are indicated. In figure B: the large black sphere indicates the zinc-ion and the small black sphere the catalytic water. All the significant residues of the active site are indicated.

Meprins are oligomeric membrane bound proteins that are abundantly expressed in the kidney tubules and small intestine of mammals. They are localised to the epithelial cell membranes where they act in the lumen of these organs. Meprins contain several additional C-terminal domains including: a MAM domain (meprin, A5 protein, receptor protein tyrosine phosphatase I), a TRAF domain (tumour necrosis factor receptor-associated factor), an epidermal growth factor-like domain (EGF), a transmembrane domain (TM) and a cytosolic tail sequence (C) (Sterchi *et al.*, 2009) (Figure 2.2). The *Tolloid*/BMP-1 proteases, which are named after the dorso-ventral patterning protein from *Drosophila* (Shimell *et al.*, 1991) and its mammalian equivalent BMP-1, are morphogenetically active proteins which are characterised by CUB (complement subcomponents Clr/Cls, embryonic sea urchin protein Uegf, BMP-1) and EGF domains (other members include SNAP and BP10) (Figure 2.2). The "hatching" enzymes, which play specific roles during embryonic development, are a group of heterogenous enzymes: some consist of only the catalytic domain, while others have additional C-terminal cysteine-rich or CUB domains (Sterchi *et al.*, 2009) (Figure 2.2).

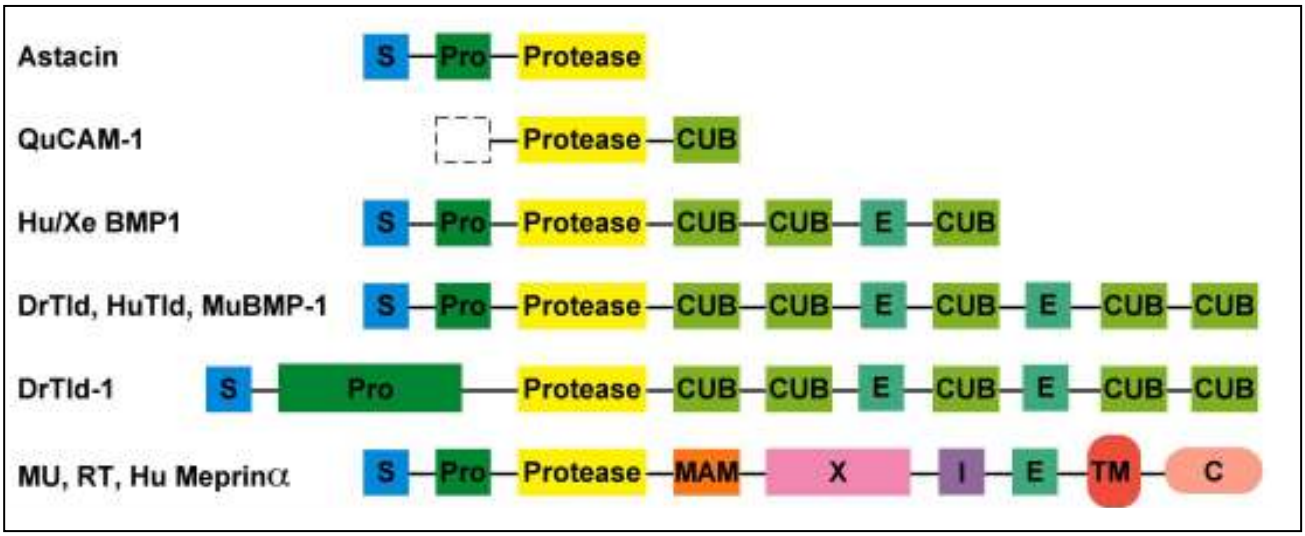

Figure 2.2. Schematic presentation of the domain structure of astacin family members (Adapted from Bond and Beynon, 1995). Abbreviations correspond to: BMP1, bone-morphogenetic protein 1; Tld: *Tolloid;* S, N-terminal signal peptides; Pro, prosequence; Protease, catalytic domains; CUB, Clr/s complement-like domains; E: EGF-like domains; MAM, adhesion domain; X, unknown domain; I, inserted domain; TM, putative transmembrane domain; C, cytoplasmic domain. Species are indicated as: Qu, quail; Hu, human; Xe, frog; Dr, *Drosophila*; Mu, mouse; Rt, rat.

2.1.1.2 Substrate specificity

The substrate specificity of astacins is primarily dependent on the structural arrangement of the segment between β-III and β-IV, which contributes to the S_1' substrate binding site. Model construction analysis, of the meprin α chain onto the structure of astacin, revealed the structural aspects that give rise to the unique substrate specificity of each astacin family member. While prototype astacin prefers an amino acid with a small aliphatic side chain as P_1' residue, other family members (such as meprin-A) accept residues with a bulky, charged side chain. In prototype astacin Pro176 strictly hinders the binding of large residues in the S_1' subsite. Whereas, in meprin-A and other astacins Pro176 is substituted by Gly176 and the following Tyr177 is deleted, resulting in the formation of a deep bulge, allowing accommodation of larger P_1' side chains (Stöcker *et al.*, 1993).

A common substrate specificity of all metzincins is that due to their remarkably long substrate binding cleft, elongated substrates (comprising of at least four amino acid residues on either sides of the scissile bond) are required for optimal cleavage. Astacin itself requires substrates of at least seven amino acids or longer (Stöcker *et al.*, 1990). Substrates of the astacin family include: ingested proteins, egg envelope, transforming growth factor-β (TGF- β) and extracellular matrix (ECM) proteins such as fibronectin and gelatin.

2.1.1.3 Regulation of astacin activity

Activation

Astacins are expressed in a tissue specific manner in mature organisms and temporally and spatially in developmental systems. For example, in crayfish astacin is synthesised in the hepatopancreas and stored extracellularly as an active protease in the stomach-like cardia, where it acts as a digestive enzyme (Bond and Beynon, 1995). Other astacins (e.g. human-BMP1, rodent and human meprins α and β and alveolin) which are involved in bone formation, embryo development and egg hatching are stored in

granules and are activated and secreted at the appropriate time (i.e. by means of regulated exocytosis) to serve their function (Bond and Beynon, 1995).

Trypsin-like proteolytic cleavage of the pro-domain constitutes a major mechanism for regulating activation (Bond and Beynon, 1995). In mature astacins, the N-termini residue (usually alanine) is salt-bridged with the family-specific glutamic acid (directly adjacent to the third histidine in the zinc-motif). This salt-bridge is critical for enzyme activity, since it promotes formation of an active enzyme. Therefore, since an elongated N-terminus (propeptide) blocks this interaction the enzyme is inactive till cleavage of the pro-domain occurs.

Inhibition

All astacins are inhibited by the broad spectrum plasma protease, α-macroglobulin. Although it is postulated that specific endogenous inhibitors play a key role in the regulation of astacin metalloproteases, natural specific endogenous inhibitors have only been identified for meprins (Sterchi *et al.*, 2009) and nephrosin (Tsai *et al.*, 2004). *In vitro*, by means of zinc-chelation and/or disulphide bond reduction, various metal-chelators (1, 10-phenanthroline, EDTA, and 2, 2'-bipyridyl) and reducing agents (cysteine, glutathione) are able to reversibly inhibit astacins. No inhibition is observed by tissue inhibitors of metalloproteases (TIMPs) or by phosphoramidon (Bond and Beynon, 1995; Barrett *et al.*, 2004).

2.1.1.4 Known biological functions of astacins

Through protein degradation, protein-protein interaction modulation, ECM turn-over, extracellular coat degradation (hatching) and highly specific cleavage events such as growth factor activation, these enzymes play essential roles in a diverse range of physiological mechanisms. Astacins are involved in digestion (Stöcker *et al.*, 1993), early embryo development (dorso-ventral patterning) (Shimell *et al.*, 1991), bone and cartilage formation (Takahara *et al.*, 1994), processing of the ECM (Sterchi *et al.*, 2009) and egg hatching (Young *et al.*, 2000).

2.1.2 The reprolysin family (MEROPS classification: clan MA(M), family M12B)

The reprolysin family is composed of 3 subgroups: the snake venom metalloproteases (SVMPs), the a Disintegrin-like and Metalloprotease (ADAMs) and the mammalian ADAMs with trombospondin motifs (ADAMTSs). To date, more than 300 members have been identified of which 34 and 19 are ADAMs and ADAMTSs, respectively (Rawlings *et al.*, 2008). The name was derived from the original members which were reptilian and reproductive metalloproteases, but these proteins are not only found in mammals and snake venoms, but also in insects, fungi and other invertebrates, including ticks (Bjarnason and Fox, 1995).

2.1.2.1 Distinguishing features of reprolysins

All reprolysins are multidomain proteins, containing at least two domains of which one is the catalytic metzincin domain (HEBGHXL/FGXXHD) (Hite *et al.*, 1992). However, most members have additional domains C-terminal to the active site, including: disintegrin domains, Cys-rich regions, C-type lectin domains, EGF-like domains, thrombospondin 1-like domains and/or transmembranes and cytoplasmic domains (Figure 2.3).

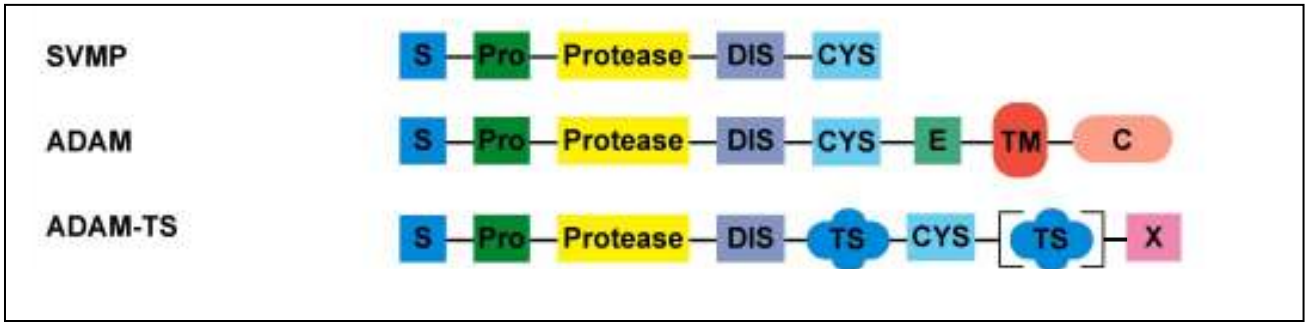

Figure 2.3. Schematic presentation of the general domain structure of the reprolysin subgroups: SVMP, ADAM and ADAMTS (Adapted from Seals and Courtneidge, 2003). Abbreviations correspond to: S, signal sequence; Pro, prodomain; Protease, catalytic domain; Dis, disintegrin domain; Cys, Cys-rich region; E, EGF-like domain; TM, transmembrane domain; C, cytoplasmic tail; TS, thrombospondin 1-like domain; X, unknown domain.)

Based on their domain structure SVMPs are further classified into four subgroups, P-I to P-IV (Hati *et al.*, 1999). The P-I class is composed of only pro- and catalytic-domains. The P-II class has an additional disintegrin domain, usually with a spacer region linking it to the adjacent catalytic domain. The P-III class has the domain construction of P-II,

with an extra Cys-rich region downstream of the disintegrin domain (Figure 2.3). Finally P-IV SVMPs contain all the domains of P-III, with an additional C-type lectin domain.

The second subgroup, the ADAMs, also have pro-, catalytic-, disintegrin-like and Cys-rich domains. In addition they also contain EGF-like domains and/or transmembrane and cytoplasmic domains (Blobel, 2005; Edwards *et al.*, 2009). Several mammalian ADAMs proteins also have thrombospondin 1-like repeats, adjacent to the disintegrin domains, these proteins comprise the third subgroup, the ADAMTSs (Porter *et al.*, 2005).

A number of SVMPs and ADAMs structures have been solved and analysed. Only the prototype structure of adamalysin II (EC 3.4.24.46), will be briefly discussed to emphasize some of the unique features of the reprolysins. The overall structure of adamalysin II exhibits an ellipsoidal shape, with a relatively shallow active site cleft. The cleft separates a large N-terminal subdomain from the smaller C-terminal domain (Figure 2.4).

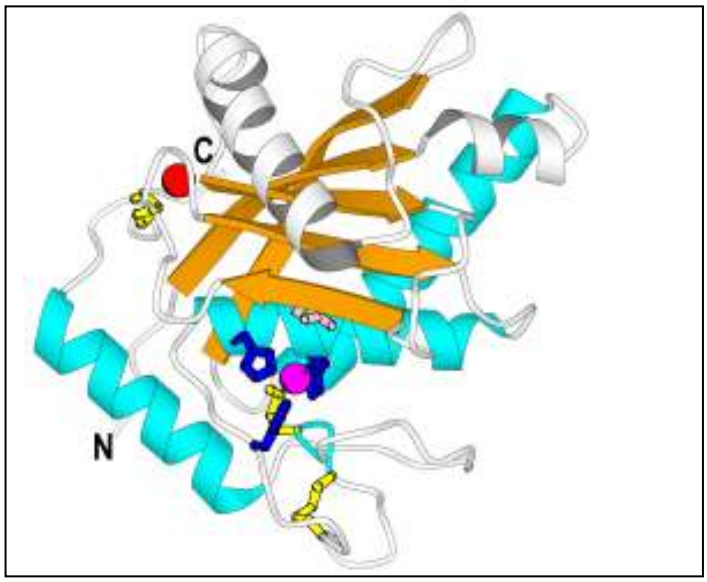

Figure 2.4. Richardson diagram of adamalysin II (PDB 1iag) (Gomis-Rüth, 2009). The common β-strands and α-helices are shown in orange and cyan, respectively. The N- and C-termini are labelled and the side chains of the zinc-binding histidines/aspartates (blue), general base (light pink), Met-turn methionine and disulphide-bonding cysteines (yellow), are displayed as sticks. The zinc ion is shown as a magenta sphere and the calcium ion as a red sphere.

In the active site cleft of adamalysin II, the catalytic zinc-ion is tetrahedral coordinated by the three conserved histidine residues and the catalytic water molecule. The family-

specific residue, following the third zinc ligand, is a strictly conserved aspartate (Asp153) that establishes a hydrogen bond with an invariant serine (Ser179), located in the first turn of α-helix C. The protein scaffold is cross-linked by two disulphide bridges connecting the N-terminal with the segment containing the Met-turn and α-helix C, respectively. Topological comparative studies of adamalysin II compared with other metzincin family lead structures revealed that reprolysins contain two additional α-helices in the N-terminal subdomain and therefore deviate from the conserved metzincin secondary structure at the β-I strand and at α-helix A and C (Gomis-Rüth *et al.*, 1993a). Another remarkable common feature of the reprolysins, also displayed in the structure of adamalysin II, is the presence of a calcium ion on the surface, located close to the C-terminal opposite the active site (Gomis-Rüth *et al.*, 1993a). Alignment of the catalytic domain of other SVMPs and ADAMs show that there are only a few single-residue insertions and deletions and all substitutions are in agreement of the lead structure. Adamalysin II exhibits 47-83% identity to other SVMPs and 27-42% identity to ADAMs (Wolfsberg *et al.*, 1995).

2.1.2.2 Substrate specificity

The cleavage specificity of reprolysins is to a large extent dependent on the S_1' substrate binding site. The S_1' subsite is a deep hydrophobic pocket, created by the loop between β-IV and β-V. Its composition results in these enzymes showing a preference for medium-sized hydrophobic uncharged residues in the P_1' position (Stöcker *et al.*, 1995; Hooper, 1996). The enzyme-substrate interactions on the C-terminal side of the scissile bond are, however, relatively similar amongst the various reprolysin members. The non-primed half of bound substrate runs anti-parallel to the upper bulge segment and parallel to the wall-forming segment below. The substrate can be linked to any of the two (or both) segments by means of hydrogen bonds (Stöcker *et al.*, 1995). Reprolysins have been found to show activity towards several substrates in the ECM including type IV collagen, nidogen, fibronectin, fibrin(ogen), laminin and gelatin.

2.1.2.3 Regulation of reprolysins

Activation

Zymogen activation provides a major regulatory mechanism and proceeds via the "cysteine switch" mechanism, which involves a cysteine residue located within the highly conserved motif, PKMCGV, of the prodomain (as has been discussed at length in section 1.5.3.1) (Springman *et al.*, 1990; Stöcker *et al.*, 1995; Gomis-Rüth, 2003).

Inhibition

Proteinaceous SVMP inhibitors include immunoglobulins found in snake venom-resistant animals such as oprin, DM40, DM43 from opossums and anti-haemorrhagic factor 1-3 from mongooses. Other SVMP inhibitors include cystatins, BJ46a and habu serum factor (Springman *et al.*, 1990; Gomis-Rüth, 2003). The regulation of the ADAMs enzymes is poorly understood and besides TIMPs, no specific endogenous inhibitors are known. Some TIMPs, such as TIMP-3 display a great selectivity towards the ADAMs proteases (Brew and Nagase, 2010). *In vitro,* reprolysins are reversibly inhibited by zinc-chelation and/or disulphide bond reduction, by metal-chelators (1, 10-phenanthroline, EDTA) or reducing agents (cysteine, glutathione), respectively.

2.1.2.4 Known biological functions of reprolysins

Differences within the primary structure of the individual domains, as well as the presence or absence of particular domains contribute to or modulate the function of reprolysin proteases. Reprolysins containing disintegrin-like domains and cysteine-rich domains are associated with haemorrhage, cell-adhesion and cell-fusion activity. EGF-like domains mediate cell-adhesion and aid cell migration. Phosphorylation of transmembrane domains and cytoplasmic tails, of ADAM- and ADAMTS-reprolysins, is known to trigger activation of intracellular signaling pathways and/or to result in the release of functional soluble ectodomains of various cell surface proteins such as growth factors, hormone receptors and cytokines (Edwards *et al.*, 2009; Reiss and Saftig, 2009). As a direct result of these multiple actions, these metalloproteases contribute to a wide variety of physiological processes. Respective functions of the

SVMP-, ADAM- and ADAMTS-reprolysins and different processes they affect are briefly discussed below.

The biological functions of SVMPs include the digestion of ECM components surrounding capillaries (like type-IV collagen, nidogen, fibrinonectin, laminin and gelatin) and inhibition of platelet aggregation (Jia *et al.*, 1996; Lu *et al.*, 2005). It is therefore evident that SVMPs contribute a great deal to snake venom-induced pathogenesis including haemorrhage, edema, hypotension, inflammation and necrosis.

ADAMs proteases were originally described from mammalian reproductive tracts, where they are vital to spermatogenesis, sperm function and fertilisation. ADAMs are also present in other somatic tissues where they regulate cell-differentiation and migration. Mutation studies revealed that they are essential in biological mechanisms such as angiogenesis, neurogenesis and myogenesis (Reiss and Saftig, 2009). These reprolysins are also involved in “ecto-domain shedding” of various cell surface proteins which lead to the activation of intracellular signaling pathways and/or the release of functional soluble ectodomains. Non-catalytic ADAMs proteins have been proven to be essential for the normal function of the central nervous system (CNS) (Edwards *et al.*, 2009). The third subgroup, the ADAMTS reprolysins, is involved in collagen processing, organogenesis, angiogenesis, embryonic development, gonad formation, inflammation and also contribute to blood coagulation haemostasis by affecting von-Willebrand factor turnover (Porter *et al.*, 2005).

2.1.3 Proposed functions of metzincin metalloproteases in ixodid ticks

Metalloproteases that form part of the array of known bioactive molecules in tick saliva have been classified as reprolysin-like metalloproteases based on sequence alignment analysis (Francischetti *et al.*, 2003). The alignment of three *Ixodes scapularis* putative reprolysin metalloproteases and atrolysins (SVMPs) indicate that the *I. scapularis* salivary gland metalloproteases contain the conserved metzincin zinc-motif and the reprolysin conserved aspartic acid residue, adjacent to the third histidine in the zinc-motif, as well as the conserved methionine (Figure 2.5). However, one remarkable difference is that the conserved proline of the Met-turn is substituted by a conserved tyrosine, giving rise to a unique Met-turn motif, GYLMSY. Furthermore, adjacent to the

active site is a cysteine-rich domain, but no RGD triplet motif - typical of the disintegrins. To date, six reprolysin-like metalloprotease genes have been identified in *Haemaphysalis longicornis* (Harnnoi *et al.*, 2007), five in *Ixodes ricinus* (Decrem *et al.*, 2007; Decrem *et al.*, 2008) and three in *I. scapularis* (Francischetti *et al.*, 2003).

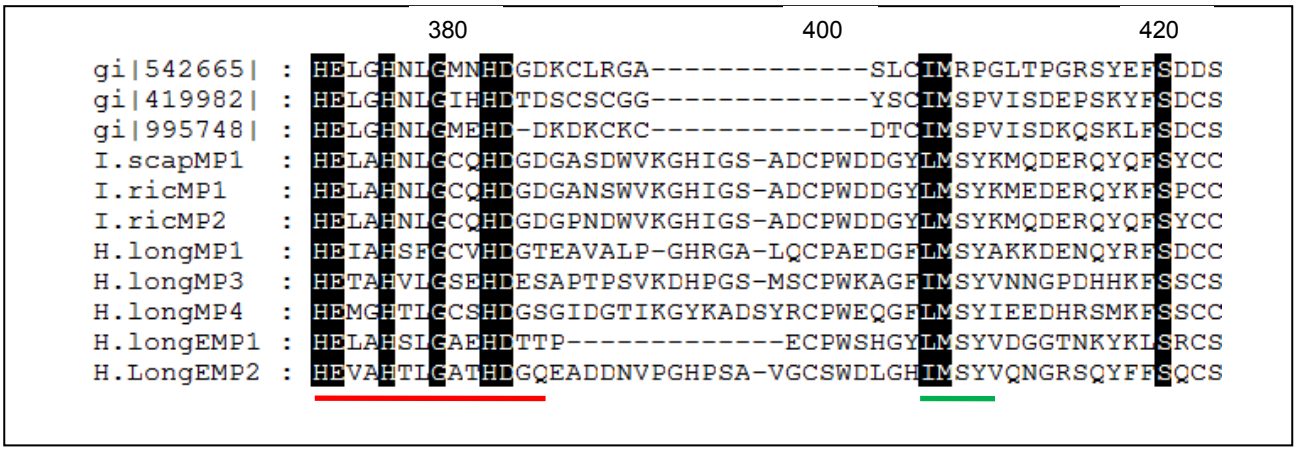

Figure 2.5. Most important features of the alignment of eight putative ixodidae reprolysins with three atrolysins, showing the conservation within the active site. (Adapted from Francischetti *et al.*, 2003). The sequences correspond to: gi|542665|, *Crotalus atrox* atrolysin B; gi|419982|, *C. atrox* atrolysin E; gi|995748|, *Protobothrops muerosquamatus* atrolysin E; I.scapMP1, *I. scapularis* MP1 (gi|60729624|); I.ricMP1, *I. ricinus* Metis1 (gi|170285591|); I.ricMP2, *I. ricinus* Metis 2 (gi|170285593|) and H.longMP1, H.longMP3, H.longMP4, H.longEMP1, H.longEMP2, five putative *H. Longicornis* metaloproteases. The position of the conserved reprolysin zinc-motif, including the family-specific aspartic acid, is indicated by the red line. The position of the conserved Met-turn is indicated by the green line. Reversed background: indicating conserved tick residues.

Based on sequence similarities between the salivary gland metalloproteases and the haemorrhagic SVMP, and preliminary activity assays (Francischetti *et al.*, 2003), it appears that these metalloproteases contribute to the facilitation and maintenance of a fluid blood feeding cavity during extended feeding periods, by performing anti-haemostatic activities such as fibrinogenolysis and fibrinolysis (Figure 2.6). Substrates of the putative reprolysin tick metalloproteases include fibrin(ogen), gelatin, and fibrinonectin but not collagen and laminin (Francischetti *et al.*, 2003). By hydrolysing fribrinogen and fibrin these enzymes prevent blood clot formation and dissolve existing clots by disaggregating the platelet and fibrin crosslink matrix.

To date, metalloproteases have not been described in any other ixodid tick tissue than the salivary glands. However various invertebrate studies have indicated metzincins, including reprolysins and astacins, in a wider variety of tissues with extensive roles in pathogen transmission (Gomis-Rüth, 2003; Barrett *et al.*, 2004), host infestation

(Williamson *et al.*, 2006) and reproduction (Bowles *et al.*, 2008). Therefore it is expected that there are metzincins that exhibit similar functions in the different ixodid tick tissues, which contributes to the tick's successful haematophagous parasitic behaviour and ultimately survival. This chapter deals with the identification of *R. microplus* astacin and reprolysin homologues and the determination of their expression profiles in the different tissue and life stages of *R. microplus.*

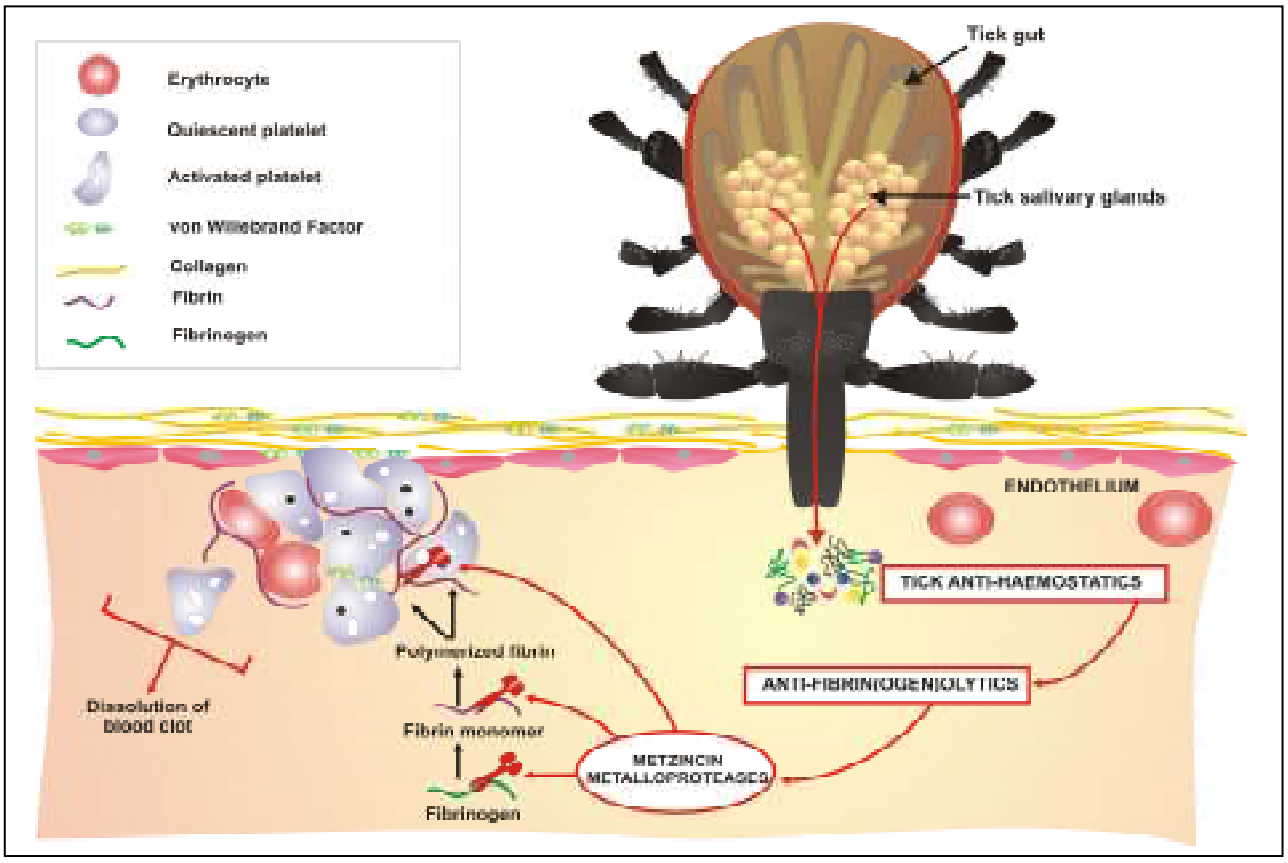

Figure 2.6. Schematic presentation illustrating the putative functions of metzincin metalloproteases present in the saliva of ixodid ticks. By hydrolysing fribrinogen and fibrin (indicated by red scissors) these enzymes prevent blood clot formation and dissolve existing clots by disaggregating the platelet and fibrin crosslink matrix.

2.1.4 Hypothesis

Reprolysin- and astacin-like metalloproteases are present in the different tissues and life stages of *R. microplus*

2.1.5 Aims

- Identification of reprolysin and astacin homologues from *R. microplus* EST databases.
- Cloning and nucleotide sequencing of reprolysin and astacin transcripts from mixed life stages cDNA by means of conventional PCR using gene specific primers.
- To determine the expression profiles of the identified reprolysin and astacin metzincins in the different tissues and life stages of *R. microplus.*

2.2 Materials and Methods

2.2.1 *R. microplus* astacin and reprolysin homologue identification

Basic Local Alignment and Search Tool (BLAST) searches were performed against the *R. microplus* Expressed Sequence Tag (EST) database of the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov), the Bmi gene index (BmiGI) of the DFCI (http://compbio.dfci.harvard.edu/tgi/) and an assembled *R. microplus* contig database (assembled by Dr. C. Maritz-Olivier). To identify astacin-like MP homologues the amino acid sequence of astacin (gi|1200203|), the family's first member (prototype), was used as query. For the reprolysin-like homologue search *Ixodes scapularis* MP1 (gi|60729624|) (Francischetti *et al.*, 2003), the first reprolysin-like MP characterised from an ixodid tick, was used as query.

Sequences with significant probabilities were identified (E-value ≤ xe^{-10}), and their mRNA nucleotide sequences were translated to their six possible reading frames using the ExPasy Translate Tool (http://ca.expasy.org/tools). The amino acid sequences were aligned by making use of Clustal W (http://www.ebi.ac.uk/clustlew/) (Thompson *et al.*, 1997) and the correct reading frames were selected, using Genedoc multiple sequence alignment editor (v 2.6.002) (http://www.psc.edu/biomed/genedoc/). Translated amino acid sequences were aligned against the amino acid sequences of astacin and *I. scapularis* MP1, respectively. Homologous sequences were selected on the presence of the conserved His-motif and conserved Met-turn residues of their metzincin family.

2.2.2 Phylogenetic analysis

For phylogenetic analysis the amino acid sequences of the identified *R. microplus* astacin and reprolysin homologues, 3 other ixodid tick reprolysin-like MPs and astacin itself were aligned using the multiple sequence alignment program MAFFT (v 6) (http://align.bmr.kyushu-u.ac.jp/mafft/software/). The information of the sequences utilised are summarised in Table 2.1. Subsequent phylogenetic analysis was performed

Table 2.1. Amino acid sequences used in phylogenetic analysis of *R. microplus* astacin and reprolysin homologues.

Name	Organism	Accession nr.	Description
BmMP1	*R. microplus*	gi|71726983|	Putative reprolysin
BmMP2	*R. microplus*	gi|71726985|	Putative reprolysin
BmMP3	*R. microplus*	gi|71726987|	Putative reprolysin
BmMP4	*R. microplus*	gi|71726989	Putative reprolysin
BmMP5	*R. microplus*	gi|71726991|	Putative reprolysin
As51	*R. microplus*	gi|82845051|	Putative astacin
As70	*R. microplus*	gi|49558770|	Putative astacin
AsContig	*R. microplus*	N/A	Putative astacin
Astacin	*A. astacus*	gi|1200203|	Astacin self
IsMP1	*I. scapularis*	gi|60729624|	Putative reprolysin
Metis1	*I. ricinus*	gi|170285591|	Putative reprolysin
Metis2	*I. ricinus*	gi|170285593|	Putative reprolysin

with the Molecular Evolution Genetics Analysis program (MEGA) (v 4.0) (Tamura *et al.*, 2007) using neighbor-joining analysis (1000 bootstraps, substitution model: PAM (Dayhoff) matrix).

2.2.3 Topology and localisation analysis

A topology analysis, in order to determine if any of the identified transcripts contain a transmembrane region, was performed using the TMHMM server (v 2.0) (http://www.cbs.dtu.dk/services/TMHMM-2.0/). To obtain a prediction of the localisation of the identified metzincins, the theoretical signal peptides and GPI-anchors were determined utilising the SignalP (v 3.0) (http://www.cbs.dtu.dk/services/SignalP/) and GPI-SOM (http://gpi.unibe.ch/) computational tools.

2.2.4 Gene specific primer design

For each identified *R. microplus* metzincin homologue a set of gene specific primers (GSPs), capable of amplifying a product of ±100 bp, was designed. The primer pair compatibility, self-complementarity and internal stability of all primer sequences were investigated with the oligonucleotide designing program Oligo version 6.0 (National Biosciences, USA) (Rychlik and Rhoads, 1989). The annealing temperature of each

primer was calculated by the following equation (Rychlik *et al.*, 1990):

Tm = 69.3 + 0.41 x (%G/C) − (650/length)

Primers were synthesised by Inqaba Biotec (Pretoria, RSA) and Invitrogen™ (The Netherlands). Primers were dissolved in 5 mM Tris-HCl (pH 7.5) and the concentration was determined spectrophotometrically (see section 2.2.9 and Table 2.5).

2.2.5 Isolation of total RNA from *R. microplus* mixed life stages

R. microplus ticks which originated from Mozambique, were provided by ClinVet International (Pty), Bloemfontein, South Africa and were subsequently bred and maintained at the Utrecht Centre for Tick-borne Disease (UCTD), University of Utrecht, the Netherlands. For cloning purposes, total RNA was isolated and purified from *R. microplus* mixed life stages (containing eggs, larvae, nymphs, adult females and adult males) by Dr. C. Maritz-Olivier. For the expression profiling studies, total RNA was isolated from 8 different *R. microplus* female tissues (haemolymph, malpighian tubules, ovaries, fatbody and both fed and unfed midgut and salivary glands), 7 different *R. microplus* male tissues (testis and both fed and unfed ancillary glands, midgut and salivary glands) and 10 different *R. microplus* life stages (eggs, unfed larvae, engorged larvae, unfed nymphs, engorged nymphs, unfed females, unfed males, partially fed females, fed males and carcasses) by Dr. A. Nijhof (UCTD).

2.2.6 Single-strand cDNA synthesis from total RNA

For conventional PCR:

To synthesise single strand cDNA from the mixed life stages RNA, a Moloney Murine Leukemia Virus (MMLV) derived reverse transcriptase (RT) with 5' tailing activity (Figure 2.7) was used. In this regard Superscript™ III RNAse H-RT (Invitrogen™ life technologies, USA) containing the modified *pol* gene (Kotewicz *et al.*, 1985; Gerard *et al.*, 1986), in conjunction with a poly-T primer (Table 2.2) was used to synthesise single-strand cDNA (100bp to >12 kb) from total RNA (specifically poly-A tailed RNA) (Figure 2.7).

Table 2.2. Primer used for mixed life stages cDNA synthesis. Degenerate nucleotide nomenclature: (N) any nucleotide and (V) any nucleotide but T.

Primer Name	Primer sequence (5' → 3')	Tm (°C)	Degeneracy
Poly-T	GCT ATC ATT ACC ACA ACA CTC T18VN	64.45	12

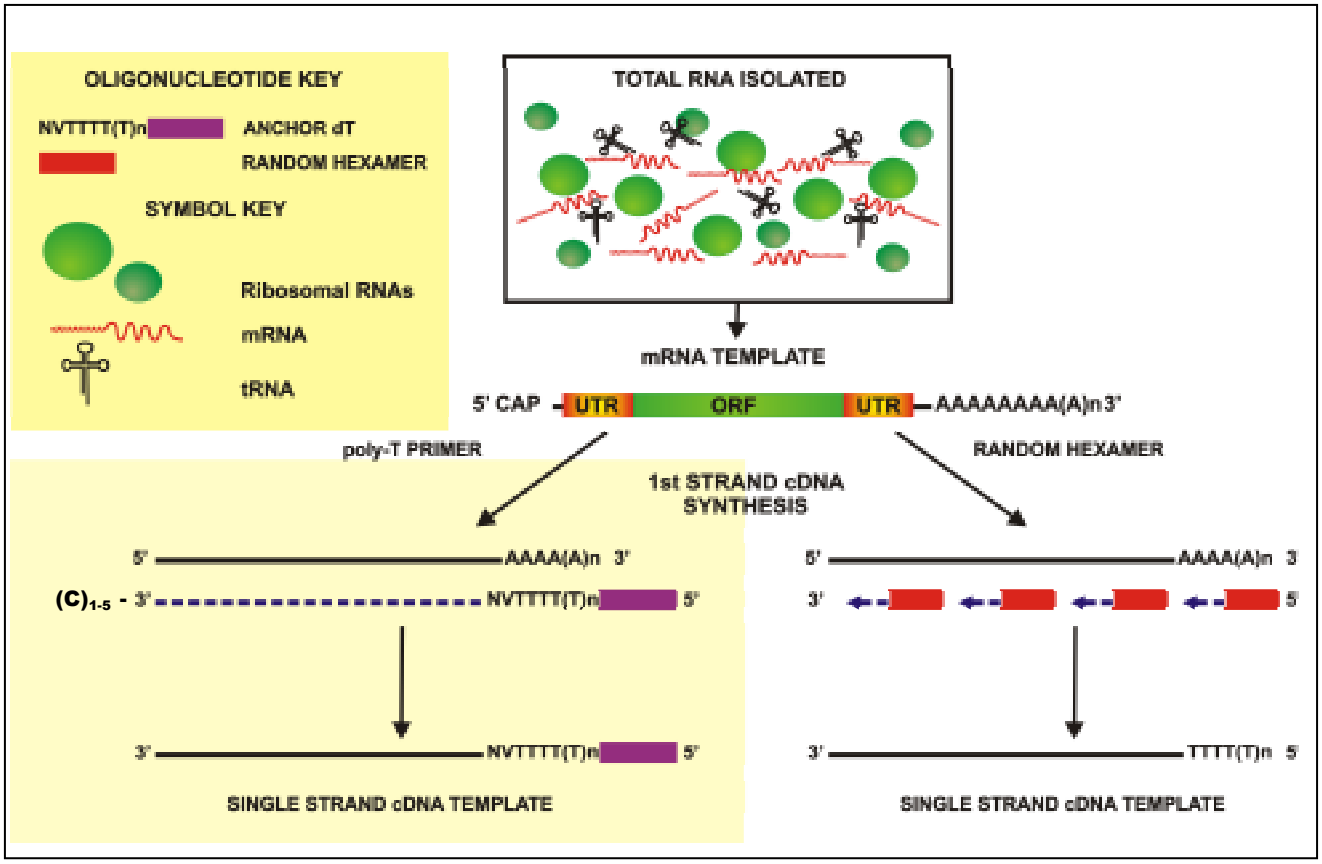

Figure 2.7. Flow diagram of single strand cDNA synthesis, using either poly-T primer or random hexamer primers. The dashed lines indicate strands being synthesised. Indicated are abbreviations for the untranslated region (UTR) and open reading frame (ORF).

Synthesis was performed according to the Superscript™ III manufacture's protocol (Invitrogen™ life technologies, USA). One microgram of total RNA was combined with 20 pmol of the poly-T primer, 0.5 mM of each dNTP (Roche Diagnostics, Germany) and nuclease-free water (Fermentas GmbH, Germany) to a final volume of 12 µl. The reaction mixture was vortexed, briefly centrifuged and the RNA was denatured by incubation at 70°C for 5 minutes. The sample was immediately chilled on ice for 5 minutes, to prevent secondary RNA structure formation, and 1x first strand buffer (250 mM Tris-HCl, 375 mM KCl, 15 mM $MgCl_2$, pH 8.3), 10 mM dithiothreitol (DTT) and 40 U of RNAse inhibitor (Promega, USA) were added. After brief centrifugation and pre-incubation at 42 °C for 2 minutes, 200 U of Superscript™ III was added. To ensure full-length synthesis, the reaction was incubated at 42 °C for 90 minutes and the

enzyme was inactivated by incubation at 70 °C for 10 minutes. Single-stranded cDNA was placed for long-term storage at -70 °C.

The 5'-anchor was incorporated, to serve as a back-up reverse primer region and to allow subsequent amplification of the entire coding sequence for future protein expression.

For real-time PCR:

To synthesise single strand cDNA from the total RNA for expression profiling and real-time PCR, the iScript™ cDNA Synthesis Kit (Bio-Rad Labrotories, Inc., USA) was used. This kit makes use of a modified MMLV-derived RT, iScript RT RNAse H+. Together with a unique blend of oligo(dT) and random hexamer primers (Figure 2.7) this enzyme yields a high quality representation of the mRNA template, including the 5'- ends. Briefly, 1 mg (for expression profiling) and 0.5 mg (for real-time PCR) of total RNA was added to, 4 µl of 5x iScript Reaction Mix, 1 µl of iScript Reverse Transcriptase and water to a final volume of 20 µl. The complete reaction mixture was incubated at 25 °C for 5 minutes, at 42 °C for a further 30 minutes, to ensure full-length synthesis. Finally the enzyme was inactivated at 85 °C for 5 minutes. Single-stranded cDNA was placed for long-term storage at -70 °C.

2.2.7 Polymerase Chain Reaction (PCR) amplification

For cloning:

All PCR reactions were performed in 200 µl thin walled tubes (Whitehead Scientific (Pty) Ltd, South Africa) and were conducted in a 2720 Thermal Cycler (Applied Biosystems, USA). Single-stranded cDNA synthesised from *R. microplus* mixed life stages was used as template. One microliter of template was combined with 10 pmol forward and 10 pmol reverse gene specific primers, 1.5 mM Mg^{2+}, 0.2 mM of each dNTP and water to a final volume of 20 µl. Two negative controls were included. These included reactions which lacked either the forward primer or reverse primer. The sample reactions were subjected to the thermal cycler in which the cDNA was denatured at 94 °C for 2 minutes (Hot-start) and then cooled to 80 °C for the addition of 5 µl enzyme

mix (1.25 U TaKaRa Taq™ (Takara, Japan) and 1x TaKaRa Taq Reaction Buffer™ diluted in water to final volume of 5 µl). Amplification consisted of 35 cycles of cDNA denaturing (94 °C, 30 sec), annealing (at the Tm of the gene specific primers, 30 sec) and extension (72 °C, 2 minutes), followed by a final extension step (72 °C, 5 minutes).

For expression profiling:

All PCR reactions were performed in 8 strip 200 µl thin wall tubes and were conducted in an iCycler Thermal Cycler (Bio-Rad Labrotories, Inc., USA). Single stranded cDNA from the different tissues and life stages was diluted 1/5 times with water and was used as template. One microlitre of template was added to a reaction solution containing 10 pmol forward and 10 pmol reverse gene specific primer and final concentrations of 1 mM Mg^{2+}, 0.1 mM of each dNTP, 1x Go*Taq* Buffer and 1.25 U of Go*Taq* (Promega, USA). The cycling reaction consisted of an initial denaturing step (94 °C, 2 minutes), followed by 35 cycles of denaturation (94 °C, 30 sec), annealing (at the Tm of the gene specific primers, 30 sec) and extension (72 °C, 30 sec), followed by a final extension step (72 °C, 7 minutes).

2.2.8 Agarose gel electrophoresis of PCR products

All PCR products were analysed on analytical graded 2% w/v agarose (Promega, USA) gels, with TAE (40 mM Tris, 1 mM EDTA, pH 8) as electrophoresis buffer at 72 V in a submarine electrophoresis unit (Amersham Biosciences, England). Each sample was loaded in 3x Blue/Orange Loading Dye (Promega, USA). All gels contained EtBr (50 µg) and DNA bands were visualised under either the Spectroline TC-312 AV transilluminator (Spectronics Corporation, USA) or an UVP imaging system (UVP, LLC., USA) at 312 nm. Relative molecular masses of the individual bands were determined with a molecular marker (Promega, USA). All PCR products were analysed in the same way.

2.2.9 Purification of PCR products

To purify the PCR products from unincorporated nucleotides, remaining polymerase, primers and salts, the pooled PCR reaction mixtures were subjected to column chromatography using the Zyppy DNA Clean and Concentrator-25™ kit (Zymo research Corporation, USA) and the High Pure PCR Cleanup Micro Kit™ (Roche Diagnostics, Germany).

The Zyppy DNA Clean and Concentrator-25™ kit (Zymo research Corporation, USA) was chosen for the purification of the ~100 bp products since it is able to purify products from 75 bp, according to the manufacturer's protocol. One volume of PCR reaction mix was added to 2 volumes of DNA binding buffer in a 1.5 ml microcentrifuge tube and vortexed briefly. The reaction mixture was loaded onto a Zymo-Spin™ column in a collection tube and was centrifuged for 30 sec at 10 000 x g. After the flow through was discarded, the column was washed twice with 200 µl wash buffer and centrifuged for 30 sec at 10 000 x g. To elute the DNA, the column was placed in a clean 1.5 ml tube and 30 µl water was loaded onto the column and the column was incubated in a 37 °C shaking incubator for 5 minutes and subsequently centrifuged at 10 000 x g for 30 sec.

The concentrations of the purified products were determined spectrophotometrically. Finally the samples were stored in low adhesion 600 µl tubes at -20 °C.

2.2.10 Nucleic acid quantification

The concentration of purified DNA was calculated from the absorbance of a 1/10 dilution at 260 nm (A_{260}), using the Gene Quant Pro Spectrophotometer (Amersham Biosciences, England) and the Gene Quant pro DNA calculator™ (Amersham Biosciences, England), which makes use of the following formula:

Concentration = A_{260} x 1 absorption unit at A_{260} x dilution factor

One absorbance unit at 260 nm is equal to 50 ng/µl double stranded DNA. The ratio between the absorbance at 260 nm and 280 nm (A_{260}/A_{280}) was used as an indication of

protein contamination and thus purity. DNA is accepted as pure when the A_{260}/A_{280} ratio is between 1.7-2.0.

Oligonucleotide (primer) concentrations were determined from their 260 nm absorbance in a 1/100 dilution using the UV-160 UV-visible recording spectrophotometer (Shimadzu Corporation, Japan). The concentrations were determined by substituting the absorbance into the Beer-Lambert equation **(A = εℓc)**:

$$\textbf{Concentration (c)} = \frac{\textbf{A}_{260}\ \textbf{(A) x dilution factor}}{\textbf{extinction coefficient (ε) x path length (ℓ)}}$$

The extinction coefficient (ε) for single stranded DNA is 33 µg/ml. A 1 cm gap cuvette was used, making the light path length (ℓ) 1 cm.

2.2.11 Cloning procedures

2.2.11.1 Ligation of the ~100 bp DNA products into pGEM®-T Easy vector

In order to sequence the ~100 bp amplified products, they were ligated into the pGEM®-T Easy vector system (Promega, USA). This vector system is used, since it allows for easy A/T cloning. The pGEM®-T Easy vector (Figure 2.8) also contains T7 and SP6 RNA polymerase promoter sites, an ampicillin resistance gene for primary selection and a gene encoding the α subunit of lacZ protein, with an internal multiple cloning site (MCS) for blue and white screening. For optimal sticky end ligation the DNA products were ligated into 50 ng of the pGEM®-T Easy vector initially in a 3:1 ratio (Sambrook *et al.*, 1989). The following formula was used to calculate the amount of DNA needed:

$$\frac{\textbf{Size of insert (Kb) x vector (ng) x (ratio of insert:vector)}}{\textbf{Size of vector (Kb)}} = \textbf{ng of insert DNA needed}$$

Ligation was performed at 16 °C overnight with 1 µl T4 DNA Ligase (3 U/µl) and 1 µl 10x Ligase buffer (Promega, USA). The reactions were heat inactivated by incubation at 70°C for 15 minutes.

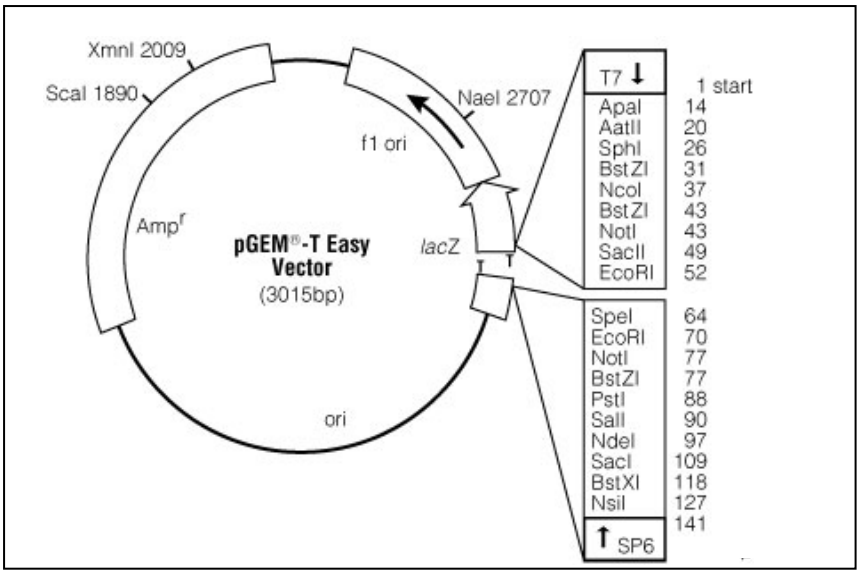

Figure 2.8. pGEM®-T Easy vector circle map

To prepare the ligation mixtures for electroporation, each reaction was precipitated with the addition of one tenth of the volume tRNA (10 mg/ml), one fifth of the volume sodium acetate (3 M, pH 5) and 3 volumes of 100% ethanol. Each reaction mixture was centrifuged for 45 minutes at 16 000 x g (4 °C) and the supernatant discarded. To remove salts, the precipitate was washed with 500 µl 70% ethanol and subsequent centrifugation. The pellet was dried *in vacuo* and dissolved in 20 µl double distilled de-ionised water and stored at -20 °C.

2.2.11.2 Preparation of electrocompetent DH5α *E. coli*

Electrocompetent DH5α *E. coli* cells (Gibco BRL, Life technologies, USA) was prepared from a single colony that was inoculated in 15 ml Luria-Berthani Broth (LB Broth, 1% NaCl, 1% tryptone, 0.5% yeast extract, in double distilled de-ionised water, ~pH 7.4-7.5) and grown overnight with moderate shaking (250 rpm) at 30 °C. Two 5 ml fractions from the overnight cultures were inoculated into two separate sterile 2-litre flasks, each containing 500 ml LB broth and were grown with shaking (250 rpm) at 37 °C to an OD_{600nm} of ~0.5 to 0.6. These cultures were equally divided into four pre-chilled 250 ml centrifuge bottles and incubated on ice at 4 °C for 20 minutes. Chilled cells were centrifuged for 20 minutes at 10 000 x g (4 °C) and the cell pellets were resuspended and washed by gentle swirling in 250 ml ice cold double distilled de-ionised water. The wash step was repeated three times. Washed cell pellets were suspended in 10 ml

10% (v/v) glycerol and pooled into two 50 ml centrifuge tubes. The samples were incubated on ice at 4 °C for 60 minutes and centrifuged to collect cells. Finally the pellets were pooled and resuspended in 1 ml 10% (v/v) glycerol. Electrocompetent cells were divided into 90 µl aliquots and stored at -70 °C.

2.2.11.3 Transformation by electroporation

Electroporation is a simple, very rapid and efficient transformation method that gives rise to plasmid transformation efficiencies of 10^8 – 10^9 cfu/µg DNA. With electroporation cells are exposed to a brief high-voltage electric pulse, this electric shock renders the bacterial cells permeable to DNA (Chassy *et al.*, 1988). Subsequently the cells take up exogenous DNA from the suspended solution and a proportion of the cells become stably transformed and can be selected.

Electrocompotent DH5α cells (90 µl) were thawed on ice and 10 µl of the ligation mixture was added to the cells and mixed gently with the pipette tip, before pipetting the 100 µl reaction mixture into a pre-chilled 0.1 cm gap Micropulser® electroporation cuvette (Bio-Rad Laboratories, Inc., USA). Electroporation was performed at 2000 V with the Electroporator 2510 (Eppendorf, Germany) for 4 milliseconds. Directly after electroporation, 100 µl LB Broth (containing 0.02 M *D*-glucose) was added to the cells in the cuvette and the cells were transferred to 900 µl LB Broth-glucose and incubated at 37 °C with shaking for 60 minutes. For each transformant three LB-agar plates (2% w/v agar in LB-Broth), containing 50 µg/ml ampicillin, were prepared onto which 50 µl of a 1/5, 1/10 and 1/50 dilution of transformed cells were plated, respectively. Plates were incubated upside down at 37 °C over night. Cells that contain the vector had ampicillin resistance and grew on the LB-ampicillin medium, whereas the cells that lack the vector were ampicillin sensitive and therefore were unable to grow on the LB-ampicillin medium.

2.2.11.4 Screening for recombinant colonies

To identify recombinant clones PCR colony screening, using the SP6 and T7 primers (Table 2.3), were performed. Eight random colonies were picked for each transformant

and used as template in a PCR mixture containing 2x KAPATaq Ready Mix DNA Polymerase (KAPABiosystems, USA) and 10 pmol each of the SP6 and T7 primers. The reaction mixture was made up to a final volume of 25 µl with water. The cycling reaction consisted of an initial denaturing step (94 °C, 7 minutes) to ensure complete cell lysis, followed by 30 cycles of denaturation (94 °C, 30 sec), annealing (55 °C, 30 sec), extension (72 °C, 2 minutes) and a final extension (72 °C, 5minutes). The PCR products were analysed on a 2% w/v agarose/TAE/EtBr gel as described in section 2.2.6.

Clones displaying the correct insert sizes were selected and re-grown overnight at 37 °C with shaking in 5 ml LB-ampicillin in 50 ml flasks. Cell stocks were made, from the overnight cultures, in LB-Broth containing 10% v/v glycerol and were stored at -70 °C.

Table 2.3. The characteristics of the T7 and SP6 primers

Primer Name	Primer sequence (5' → 3')	Tm (°C)
T7	GTA ATA GCA CTC ACT ATA GGC GA	58.87
SP6	ATT TAG GTG ACA CTA TAG AAT AC	53.52

2.2.11.5 Plasmid isolation using the Zyppy™ Miniprep Kit

Pure plasmid for further analysis was isolated using the Zyppy™ Miniprep Kit (Zymo research Corporation, USA). This isolation kit makes use of alkaline lysis to release DNA from the cells. The alkaline solution is neutralised with a high salt, low pH buffer which establishes conditions that favor plasmid re-folding. The neutralisation buffer also contains RNAse A which removes contaminating RNA. Due to its size genomic DNA does not anneal fully and thus can be removed together with the proteins and other cell debris by centrifugation.

Cells from 3 ml of the overnight bacteria culture were collected by centrifugation (13 000 x g, 1 minute), resuspended in 600 µl water, 100 µl of 7x Lysis Buffer and mixed by inverting the tube 4-6 times. Cells were incubated for 2 minutes at room temperature to allow complete lysis before 350 µl of cold Neutralization Buffer was added and the sample was mixed thoroughly by tube inversion. Genomic DNA and cell debris was removed with centrifugation (13 000 x g, 4 minutes) and the plasmid-containing supernatant transferred to a Zymo-Spin™ column. The column was placed in

a collection tube, centrifuged (13 000 x g, 15 sec) and the flow-through discarded. The column was washed twice, once with 200 µl endo-wash buffer and once with 400 µl Zyppy™ Wash Buffer and centrifugation (13 000 x g, 30 sec). Finally, the column was transferred into a clean 1.5 ml microcentrifuge tube, 30 µl of elution buffer added, incubated at 37 °C with shaking for 1 minute, and subsequently centrifuged (13 000 x g, 15 sec). The concentration of the purified plasmid DNA was determined spectrophotometrically and samples were stored at -20 °C.

2.2.12 Automated nucleotide sequencing and data analysis

DNA nucleotide sequences of recombinant clones were determined by automated nucleotide sequencing using the ABI PRISM® DT3100 Genetic Analyzer (PE Applied Biosystems, USA) and the Big Dye version 3.1 sequencing kit (Perkin Elmer, USA). Automated sequencing makes use of dideoxynucleotide chain terminators (nucleotides which lack the 3' hydroxyl group), which upon incorporation cause early chain termination at different stages of strand synthesis. Each of the four nucleotides is coupled to a different fluorophore, with a characteristic excitation and emission wavelength, to determine the sequence in a single sequence run.

Each sequencing reaction contained 3 µl 5x sequencing buffer (400 M Tris-HCl, 10 mM $MgCl_2$, pH 9), 2 µl V3.1 Big Dye (Perkin Elmer, Foster City), 1 µl T7/SP6 primer (5 pmol), 550 ng plasmid and water to a final volume of 20 µl. Cycle sequencing was performed using 25 cycles, which included denaturation (94 °C, 30 sec), annealing (50 °C, 30 sec) and extension (60 °C, 4 minutes).

Ethanol precipitation was conducted to purify the amplified DNA from excess primers/salts and unincorporated fluorescent ddNTP's. To each sample one fifth of the volume sodium acetate (3 M, pH 5) and 3 volumes of 100% ethanol were added and the reaction mixture was centrifuged for 45 minutes at 16 000 x g (4 °C). The supernatant was discarded and the precipitate washed with 500 µl 70% ethanol and subsequent centrifugation. The wash step was repeated twice. The pellet was dried *in vacuo* in the vacuum concentrator (Bachoffer, Germany) and analysed according to the protocol outlined in the ABI PRISM® DT3100 Genetic Analyzer user's manual

(PE Applied Biosystems, USA), by the DNA Sequencing Facility of the Faculty of Natural and Agricultural Sciences at University of Pretoria (http://seqserve.bi.up.ac.za).

DNA sequences obtained were analysed using the BioEdit version 5.0.9 program (Hall, 1999). Once the results were confirmed and correlated by visual inspection of the electropherograms the nucleotide sequences were aligned against their original astacin or reprolysin nucleotide sequences using Clustal W (http://www.ebi.ac.uk/clustlew/) (Thompson *et al.*, 1997).

2.3 Results and Discussion

2.3.1 *R. microplus* astacin- and reprolysin-like homologue identification

For the reprolysins, EST sequences from *R. microplus* with the highest similarity (E-values ≤ xe^{-10}) to *I. scapularis* MP1 were identified and aligned. Sequences that did not possess the conserved reprolysin His-motif and Met-turn were discarded. This resulted in the identification of several sequences from *R. microplus* that may form part of the reprolysin family of metzincins. For five of the sequences the complete coding sequence (cds) was available and these were selected for further studies: BmMP1: gi|71726992|, BmMP2: gi|71726990|, BmMP3: gi|71726988|, BmMP4: gi|71726986| and BmMP5: gi|7172684|. For sequence analysis the amino acid sequences of the five reprolysin-like *R. microplus* homologues were aligned with 2 reprolysin SVMPs (Figure 2.9) and *I. scapularis* MP1 (Figure 2.10).

Close inspection of the alignments reveal that all the *R. microplus* reprolysin-like sequences contain the typical zinc-binding (catalytic) domain of the reprolysin family (HEBGHXL/FGXXHD), except for one conserved glycine to alanine substitution in BmMP1, BmMP2, BmMP3 and BmMP5. In BmMP4 the third residue of the motif is substituted with a cysteine, instead of the usual hydrophobic amino acid. The alignment against known reprolysins indicates the conserved methionine adjacent to the catalytic

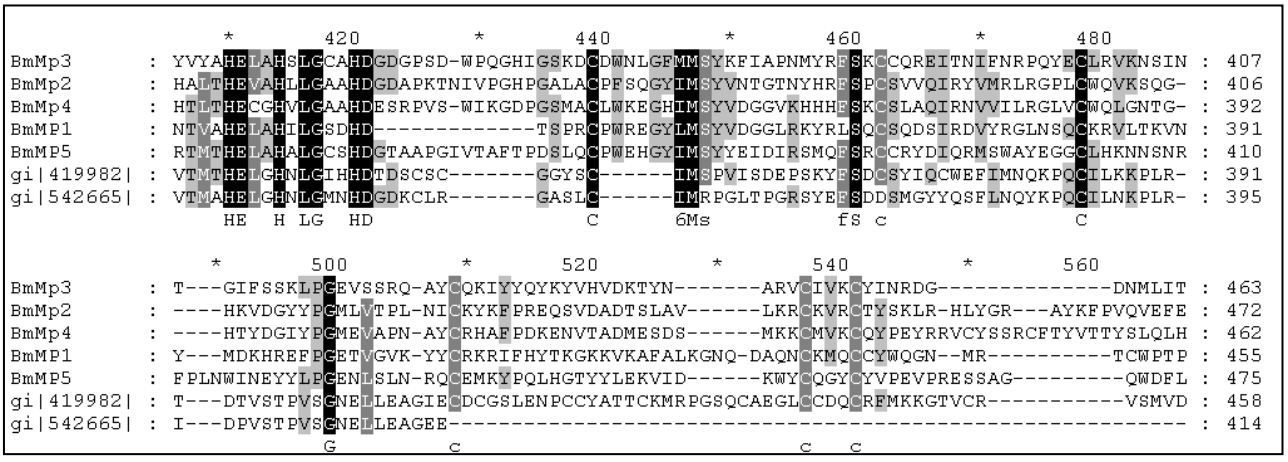

Figure 2.9. Amino acid alignment of the five *R. microplus* putative reprolysins against two reprolysin SVMP, atrolysins. Sequence similarity is represented by grey scaling, with black indicating 100% similarity, grey 80-90% similarity and light grey 60-70% similarity. The atrolysin sequences correspond to: gi|542665|, *Crotalus atrox* atrolysin B: gi|419982|, *C. atrox* atrolysin E.

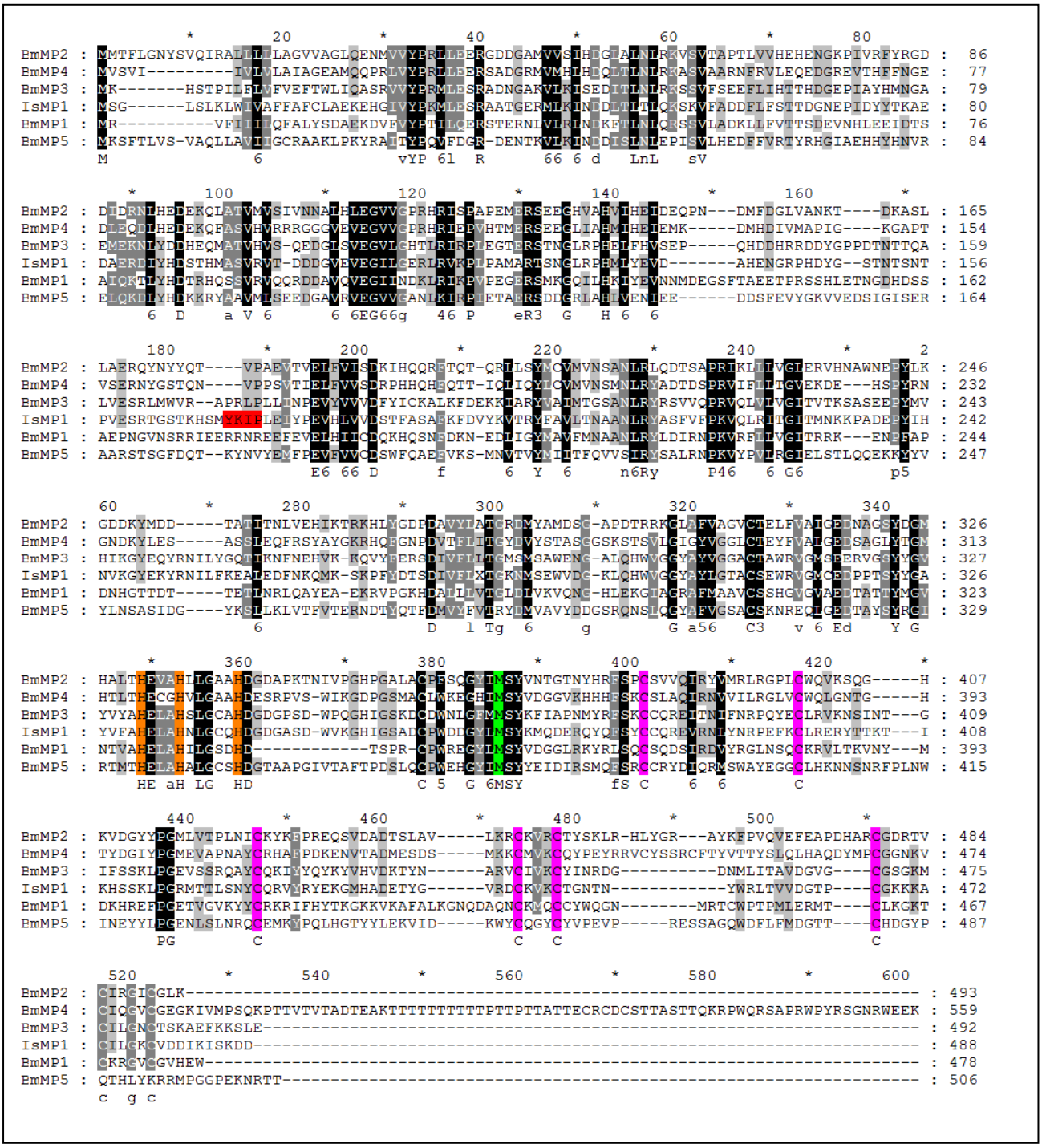

Figure 2.10. The amino acid alignment of the 5 reprolysin-like *R. microplus* sequences against *I. scapularis* MP1 (IsMP1, gi|60729624|). Sequence similarity is represented by grey scaling, with black indicating 100% similarity, grey 80-90% similarity and light grey 60-70% similarity. The mature IsMP1 protein starts with the YKIP peptide (shown in red reverse background) obtained by Edman degradation of salivary proteins (Valenzuela *et al.*, 2002). The catalytic domain conserved histidine is highlighted in orange and the conserved methionine in green. All cysteines are shown in magenta reversed background.

domain (Figure 2.10). For the ixodid tick reprolysin-like sequences a unique conserved Met-turn motif is observed: GXXMSY. This Met-turn is similar to the conserved Met-turn of the astacin family, in which the second residue after the methionine is a conserved tyrosine (Gomis-Rüth, 2003). It has been postulated, due to its aromatic structure, that the conserved tyrosine may play a role in substrate binding and/or zinc-coordination (Park and Ming, 1998). Interestingly, it has been found that metzincins which contain this conserved tyrosine lack the "Cys-switch" motif for activation, most probably since the phenolic side chain protects the metal in the unbound state. All 5 of the *R. microplus* reprolysin-like MPs do in fact lack the "Cys-switch" motif, PKMCGV. Therefore it can be reasoned that these putative reprolysins are most likely activated by trypsin-like proteolytic cleavage, rather than the "Cys-switch". None of the *R. microplus* sequences contain the conserved RGD triplet typical of the disintegrin-domains present in some reprolysins. Finally, from the alignment it is evident that each of the 5 *R. microplus* reprolysin-like MPs possess a Cys-rich C-terminal domain. Although the function of these conserved cysteines are unknown, it is speculated that similar to other MPs they provide binding sites to ECM proteins (Iba *et al.*, 1999).

For the astacins, the EST sequences with the highest similarity (E-values ≤ xe^{-10}) to prototype astacin were identified and aligned. The sequences that did not possess the conserved astacin His-motif and Met-turn were discarded. This resulted in the identification of several EST sequences from *R. microplus* that may form part of the astacin family. However with further data analysis it was detected that more than one EST sequence represented the same gene. BLAST searches were therefore conducted with all the obtained EST sequences against an assembled *R. microplus* contig database. From the integrated BLAST searches and alignments (Figure 2.11) three possible astacin-like *R. microplus* homologues were identified; As51: gi|82845051|, As70: gi|49558770| and AsContig.

From the alignment of the 3 *R. microplus* astacin-like MPs against the original crayfish astacin and another arthropod astacin-like MP (a toxin precursor of *Loxosceles intermedia*), it is evident that all tick MPs contain the typical zinc-binding (catalytic) domain of the astacin family (HEBXHA/VBGFXHE), except for one glutamic acid to methionine substitution in As70. The alignment also indicates that the conserved Met-turn is present in all 3 *R. microplus* astacin-like MPs, however, the histidine directly

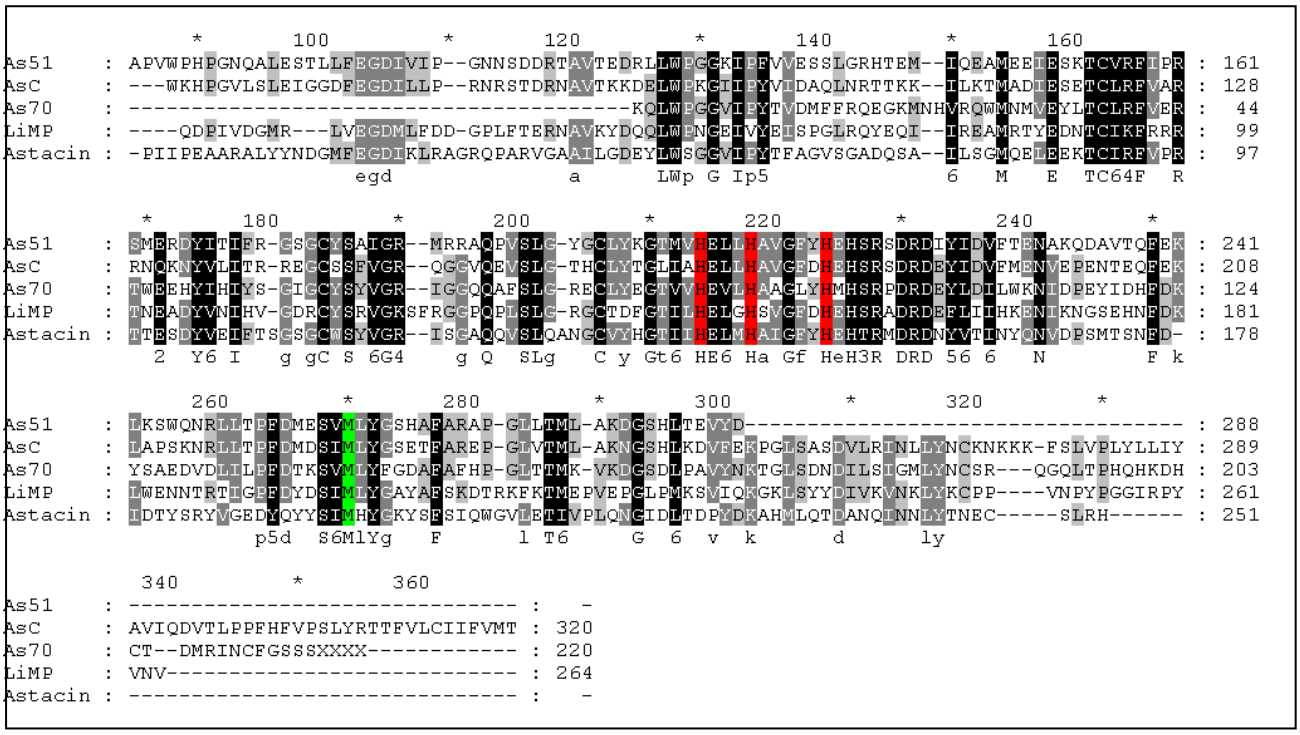

Figure 2.11. Amino acid alignment of the 3 astacin-like *R. microplus* sequences against prototype astacin (gi|1200203|) and astacin-like MP toxin precursor of *Loxosceles intermedia* (LiMP, gi|116733934|). Sequence similarity is represented by grey scaling, with black indicating 100% similarity, grey 80-90% similarity and light grey 60-70% similarity. The catalytic domain conserved histidines are highlighted in red and the conserved methionine in green.

adjacent to the methionine is substituted with a lysine, but the structurally important tyrosine remains conserved. The typical astacin connecting segment (RXDRD), between the His-motif and the Met-turn is also present in all 3 *R. microplus* sequences. Unfortunately to date the full length sequences remain unknown, hindering further sequence analysis. As with all members of the metzincins the overall similarity and identity among the reprolysins and astacins were low, however the respective family-specific conserved motifs were present.

2.3.2 Phylogenetic analysis of *R. microplus* reprolysin and astacin homologues

Phylogenetic analysis revealed the distinct evolutionary patterns of the two different metzincin families. In the reprolysins clan, BmMP3 was closest related to the *Ixodes* reprolysin MPs. Except for BmMP5 which branches independently, the reprolysin-like *R. microplus* MPs, BmMP1, BmMP2, BmMP3 and BmMP4, are closely related to one another (Figure 2.12). As expected, the astacins grouped independently of the

reprolysins. Although the bootstrap values of the astacin clan are not ideal, it is clear that the 3 *R. microplus* astacin-like MPs are homologues to astacin, the prototype of the family (Figure 2.12).

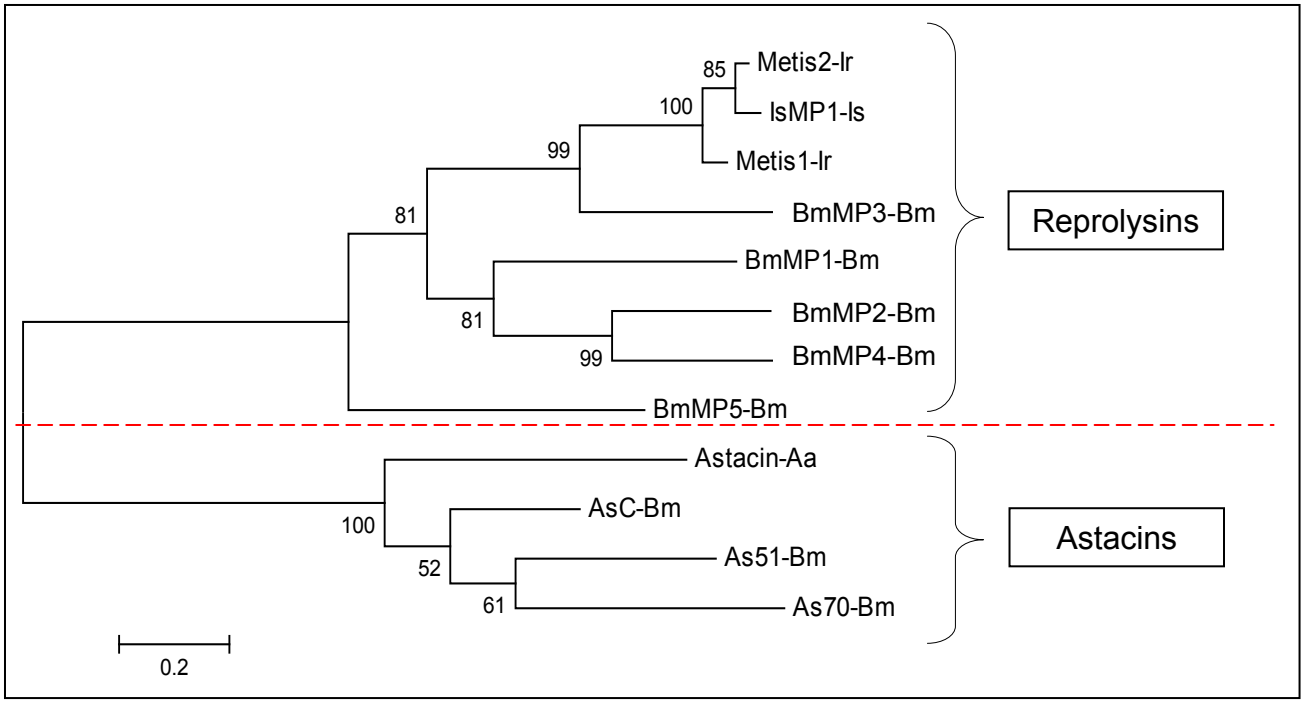

Figure 2.12. Phylogenetic analysis of the 5 reprolysin-like and 3 astacin-like *R. microplus* MPs together with other ixodid tick MPs and astacin. See Table 2.1 for all sequence information. Bootstrap values are shown on the lineage of the tree.

2.3.3 Topology and localisation analysis

As expected all 5 full length *R. microplus* reprolysin-like sequences contained a secretory signal peptide, indicating that these proteases are most likely secreted and act extracellularly (Table 2.4). For the 3 *R. microplus* astacin-like MPs the entire coding sequences were not available and therefore the presence of signal peptides in the N-terminal ends of these MPs could not be determined. However, to date all identified astacins have been proven to contain a prepro-segment (Bond and Beynon, 1995; Gomis-Rüth, 2003; Sterchi *et al.*, 2009). The presence of these prepro-domains, which direct the proteins into the ER during biosynthesis, are therefore an indication that all astacins are either secreted or transmembrane bound (Sterchi *et al.*, 2009).

To determine if any of the identified metzincins are possibly membrane-associated a topology and GPI-anchor analysis was performed. The results confirmed that all 5

reprolysin-like sequences are most likely secreted, since no transmembrane α-helices or GPI-anchors were identified (Table 2.4). Although no GPI-anchors were identified in any of the astacin sequences, it cannot be concluded that these proteins do not contain these transmembrane associated domains, since the entire coding sequences remain unavailable. Within the available sequence data no transmembrane α-helices were identified (Table 2.4). From literature it is evident that members of the astacin family present in invertebrates, which are involved in food digestion and egg hatching, lack transmembrane domains, and upon regulated secretion are released extracellularly where they function (Sterchi *et al.*, 2009). Based on the preliminary results and literature it can be speculated that the three *R. microplus* astacins are most likely secreted.

Table 2.4. Computational determination of predicted signal peptides and GPI-anchoring of *R. microplus* metzincins. Indicated are signal peptide predictions done with SignalP (http://www.cbs.dtu.dk/services/SignalP/), GPI-anchoring predictions with GPI-SOM (http://gpi.unibe.ch/) and transmembrane (TM) α-helices analysis with the TMHMM server (http://www.cbs.dtu.dk/services/TMHMM-2.0/).

	Signal Peptide (SignalP)	Potential GPI-anchor (GPI-SOM)	Potential TM α-helix (TMHMM)
BmMP1	Yes	None	None
BmMP2	Yes	None	None
BmMP3	Yes	None	None
BmMP4	Yes	None	None
BmMP5	Yes	None	None
As51	Unknown	Unknown	None
As70	Unknown	Unknown	None
AsContig	Unknown	Unknown	None

2.3.4 Gene specific primer design and PCR amplification of *R. microplus* reprolysin and astacin homologues

A set of gene specific primers (GSPs) (forward and reverse) was designed for each of the 5 reprolysin-like and 3 astacin-like *R. microplus* MPs. The primers were selected randomly to amplify a 100 bp fragment for each transcript investigated. The latter was performed to assess sequence conservation between the Winkel and Mozambican

strains that were used for EST database sequences and tissue collection, respectively. The characteristics of these primers are summarised in Table 2.5.

The latter primers and single stranded cDNA, synthesised from *R. microplus* mixed life stages total RNA was used to amplify 100 bp fragments of the reprolysin-like and astacin-like *R. microplus* transcripts. Upon analysis of the products on a 2% Agarose/TAE/EtBr gel it was clear that for all transcripts a single distinct band, of approximately 100 bp, was obtained (Figure 2.13). The negative controls showed no background amplification.

Large-scale PCR reactions were performed under the same conditions for subsequent amplification of sufficient amounts of product for cloning. The products were purified directly from the PCR reactions and the concentrations determined. Each of the 5 reprolysin-like and 3 astacin-like 100 bp fragments were cloned into pGEM®-T Easy vector and transformed into DH5α *E. coli.*

Table 2.5. Characteristics of the gene specific primers used for PCR amplification of the 5 reprolysin-like and 3 astacin-like transcripts.

Primer Name	Primer sequence (5' → 3')	Tm (°C)
BmMP1fw	GCA AGA GAG ATC AAC CGA AAG	60.61
BmMP1rv	AAA GAG AAG TTT GTC CGC AAG TA	59.20
BmMP2fw	AAT CCG TGC GCT ATT GTT GCT AC	62.77
BmMP2rv	GCC AAT CCA TCG TGA ATG CTA AC	62.77
BmMP3fw	TAA GTG AAG ACA TAA CGC TGA AC	59.20
BmMP3rv	GCT CGT GGT CGT CGT AAA	59.90
BmMP4fw	CGG GAC ACA CTT ACG ACG GAA TA	64.55
BmMP4rv	TCC ATG TCA GCG GTC ACG TT	62.45
BmMP5fw	ACG GAA CGA AAT GAC ACA TAT CA	59.20
BmMP5rv	GCA GAT CCA ACA AAG GCA TAA C	60.81
As5051fw	AAC CGA CTG CTG ACG CCA TT	62.45
As5051rv	CTC TGT CAA GTG ACT GCC GTC CT	66.33
As8770fw	GCT GAC CTC CAA TGC GAC C	64.48
As8770rv	TAC TTG ACC TGC CTG CGA TTT GT	62.77
AsContigfw	TTG AGG TGA CTG CCA TTC TTC GC	64.55
AsContigrv	AGA ACA GGT TGC TGA CGC CCT TC	66.33

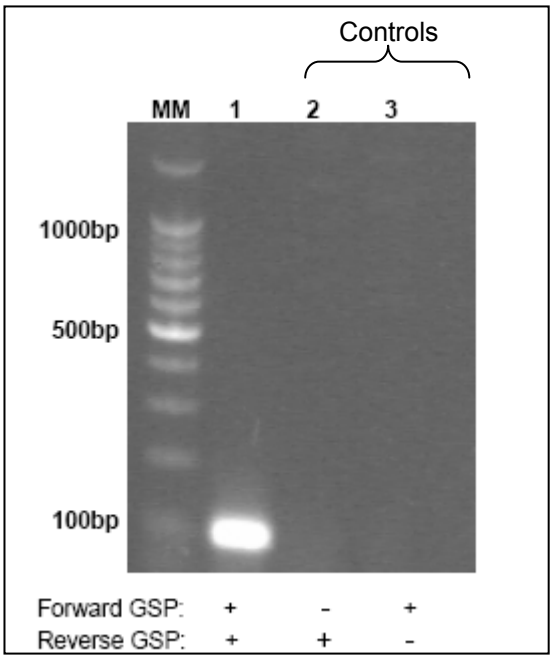

Figure 2.13. A representative sample of PCR amplification of a 100 bp fragment of a *R. microplus* metzincin, using GSPs and cDNA created from *R. microplus* mixed life stages. Amplification conditions indicated below. MM represents the 100 bp molecular marker.

2.3.5 Screening of cDNA inserts by colony PCR and gel electrophoresis

Recombinant clones were identified by colony PCR screening, using the T7 and SP6 primers (Table 2.3). These primers anneal to the pGEM®-T Easy vector at locations just outside of the multiple cloning site (Figure 2.8), which implies that all PCR products contain an additional 141 bp vector-derived sequence. For each transcript 8 colonies were screened by PCR amplification and gel electrophoresis. A number of positive clones, which were identified by distinct bands of approximately 250 bp, were obtained for each transcript (results not shown).

2.3.6 Identification of cDNA inserts using nucleic acid sequencing

In order to determine the nucleic acid sequences of the reprolysin- and astacin-like 100 bp cDNA inserts the positive clones were submitted for automated nucleic acid sequencing. For each transcript, the obtained sequences were aligned with their original EST/cds sequences (Figure 2.14).

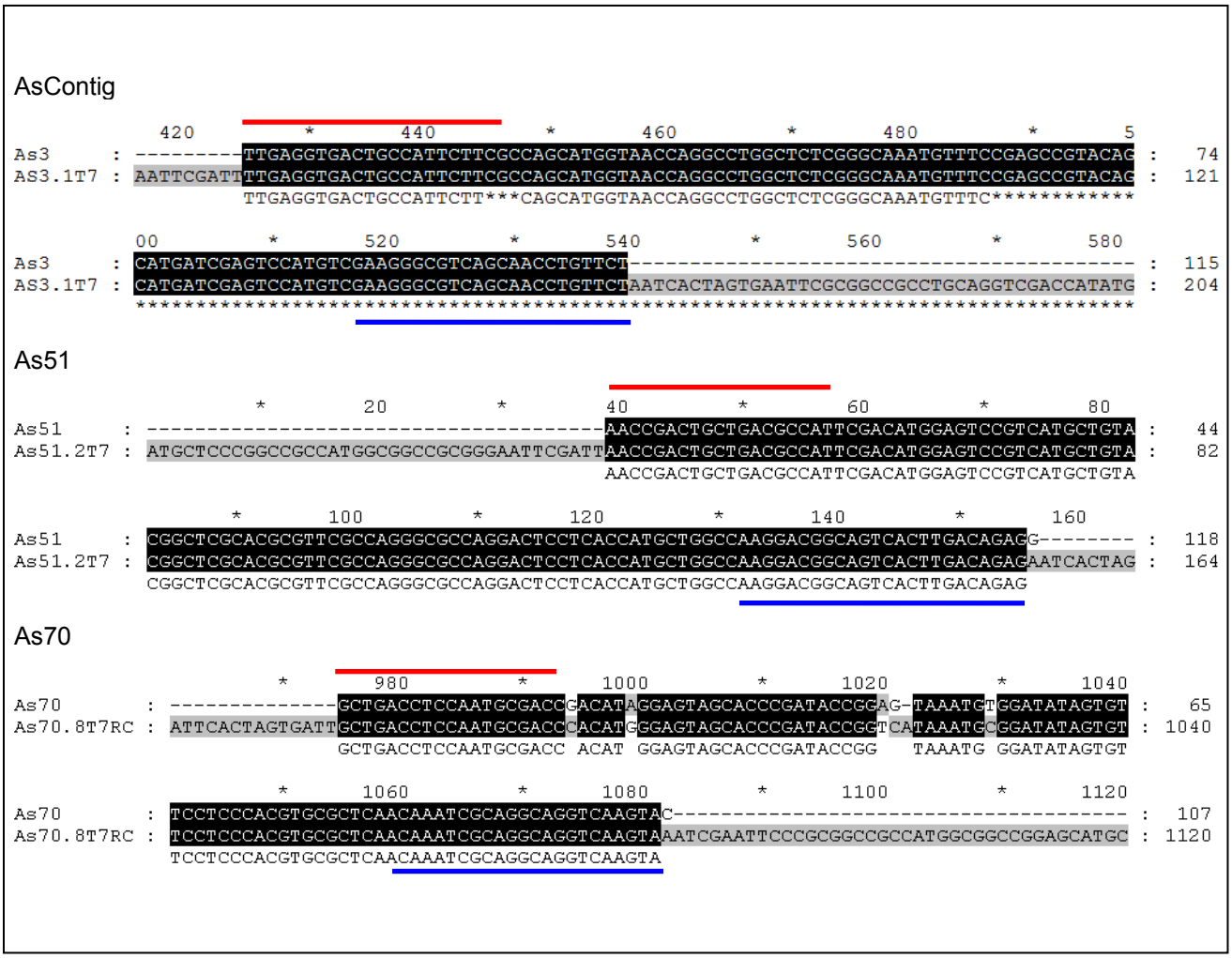

Figure 2.14. Nucleic acid sequence alignments of the 100 bp amplified fragments of the 3 astacin-like transcripts compared to their original EST sequences. Black reverse background indicates 100% similarity. For each sequence the forward and reverse primers are indicated with a red and blue line, respectively.

Results indicate that transcripts BmMP1, BmMP3, BmMP4, AsContig and As51 were successfully amplified using gene specific primers. A single basepair mutation was observed in the sequence of BmMP2. This mutation, however, does not affect the translated protein sequence, since the same amino acid (valine) is encoded for. Within the reverse primer region of BmMP5 several mutations, including a gap mutation, were observed. In the sequence of As70 there were also a few mutations, including the incorporation of an adenine. Although significant mutations were identified, the overall sequence identity, of all eight transcripts, is sufficient for future dsRNA synthesis and subsequent RNAi studies.

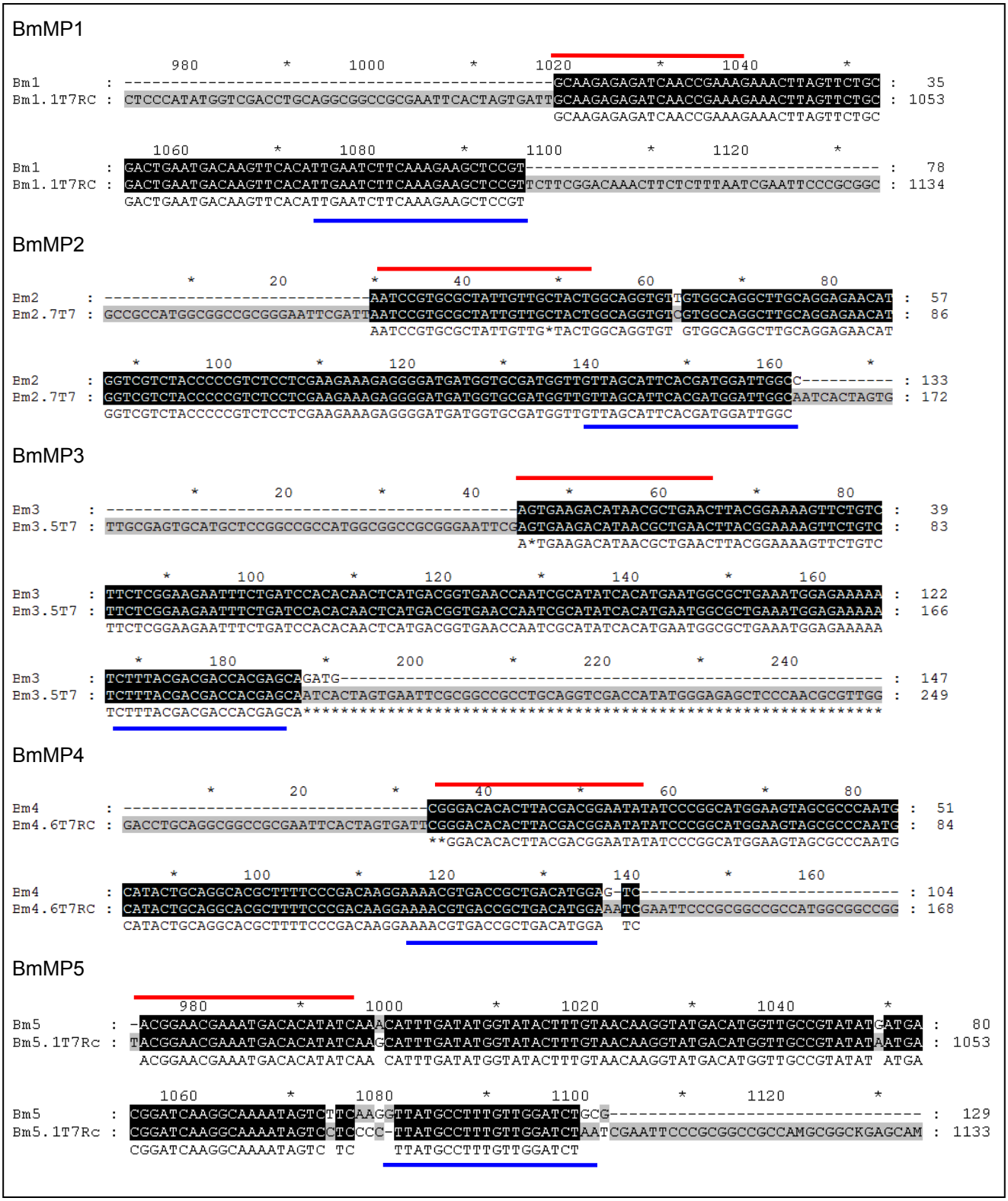

Figure 2.15. Nucleic acid sequence alignments of the 100 bp amplified fragments of the 5 reprolysin-like transcripts compared to their original coding sequences. Black reverse background indicates 100% similarity. For each sequence the forward and reverse primers are indicated with a red and blue line, respectively.

2.3.7 Expression profiling by PCR amplification and gel electrophoresis

Conventional RT-PCR was used to determine the expression profile of the 5 reprolysin-like and 3 astacin-like *R. microplus* MPs. The gene specific primers (Table 2.5) and ss cDNA, synthesised from 8 different female tissues, 7 different male tissues and 10 different feeding life stages of *R. microplus* were used. All products were visualised on 2% Agrose/TAE/EtBr gels (Figure 2.16 and 2.17).

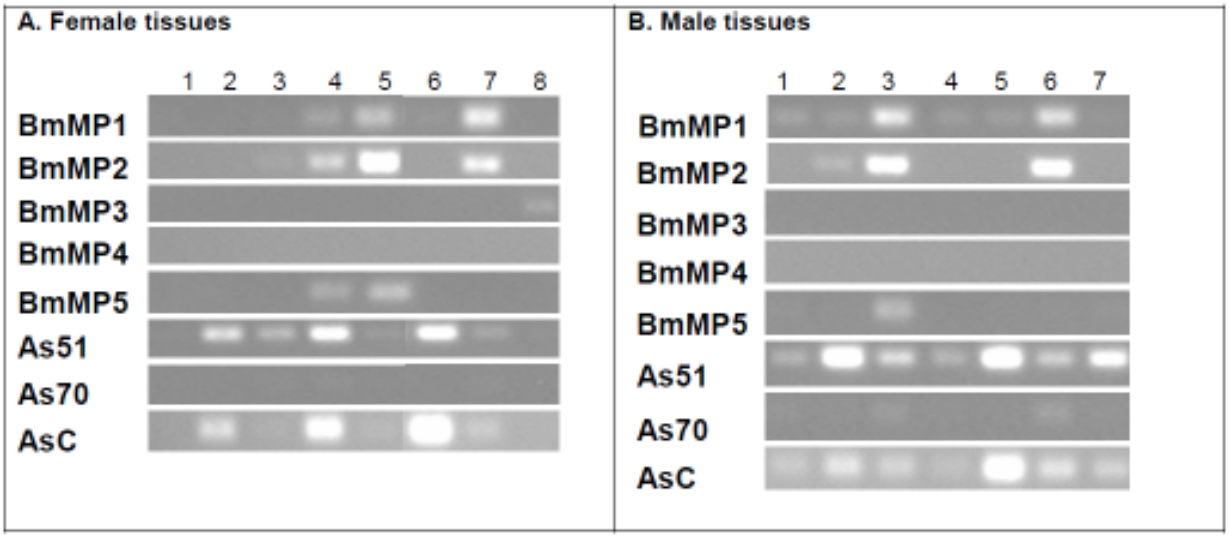

Figure 2.16. Expression profiling of reprolysin- and astacin-like metzincins in *R. microplus* adult tissues. Figure A: Female tissues. Lane 1: engorged female haemolymph, lane 2: partially fed female midgut, lane 3: partially fed female malpighian tubules, lane 4: partially fed female ovaries, lane 5: partially fed female salivary glands, lane 6: unfed female midgut, lane 7: unfed female salivary glands, lane 8: partially fed female fatbody. Figure B: Male tissues. Lane 1: fed male ancillary glands, lane 2: fed male midgut, lane 3: fed male salivary glands, lane 4: unfed male ancillary glands, lane 5: unfed male midgut, lane 6: unfed male salivary glands and lane 7: unfed male testis.

Analysis of the tissue profiles (Figure 2.16) indicated that the reprolysin-like transcripts are most abundantly expressed in the salivary glands. This is in agreement with previous reported studies of other ixodid tick reprolysin-like MPs (Francischetti *et al.*, 2003; Decrem *et al.*, 2007; Harnnoi *et al.*, 2007). These results support the hypothesis that these MPs are most likely active in the feeding pool, where they proteolytically act on host proteins to establish and maintain the fluid feeding cavity. Although RT-PCR is only semi-quantitative and further quantitative data is needed to draw final conclusions, it is interesting to note that the *R. microplus* reprolysin-like MPs displayed different expression patterns during feeding. Based on the intensity of the bands it can be

observed that BmMP1 was markedly down regulated in the female salivary glands upon feeding. In contrast BmMP2 and BmMP5 were induced with feeding. It is of great interest that the reprolysin-like transcripts were not only detected in the salivary glands but also in other adult tissues. Although it was only at low levels, BmMP1, 2 and 5 were also detected in the ovaries and BmMP3 was detected in the fatbody.

From the astacin tissue expression profiles it is evident that As51 and AsContig are abundantly expressed in both the female and male midgut. Based on the presence in the midgut and previous studies of other invertebrate astacin-like MPs (Bond and Beynon, 1995), it can be reasoned that these tick astacin homologues are most probably involved in bloodmeal digestion. Since both As51 and AsContig were detected at high levels in the reproductive organs (ovaries and testis) we hypothesise that these metzincins may be involved in tick reproduction. In the ovaries and testis these putative astacins may be involved in oogenesis and spermatogenesis, respectively. Both As51 and AsContig were not only detected in midgut and reproductive organs but also in lower levels in the salivary glands, malpighian tubules and ancillary glands. This is the first report of reprolysin-like and astacin-like metalloproteases in tissues other than the salivary glands of ixodid ticks.

Analysis of the life stage profiles (Figure 2.17) indicates that As51 is the most abundantly expressed transcript throughout all life stages. It is the only transcript

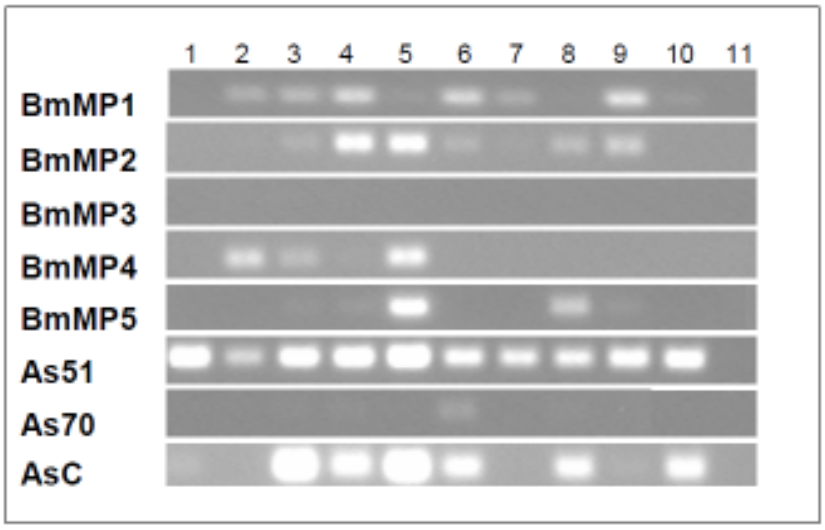

Figure 2.17. Expression profiling of metzincins in *R.microplus* feeding life stages. Lane 1: eggs, lane 2: unfed larvae, lane 3: engorged larvae, lane 4: unfed nymphs, lane 5: engorged nymphs, lane 6: unfed female, lane 7: unfed male, lane 8: partially fed female, lane 9: fed male, lane 10: carcass and lane 11: no template control.

detected in eggs. Since astacins in other invertebrates are involved in embryogenesis or egg hatching (Bond and Beynon, 1995), it can be hypothesised that this *R. microplus* astacin may have a similar function. The expression profiles of BmMP1, 2 and 5 in the different feeding life stages followed the same trend as their tissue profiles: BmMP1 was down regulated and BmMP2 and BmMP5 was up regulated with feeding. Although BmMP3 could not be detected initially, upon amplification with newly synthesised primers, it was detected in fed nymphs, partially fed adult females and mixed adults. BmMP4 is only detected in the immature stages.

2.4 Conclusion

BLAST searches of the NCBI and an assembled contig database of *R. microplus* ESTs yielded five sequences with significant similarity to a known reprolysin-like ixodid tick MP, and three sequences with significant similarity to prototype astacin. Analysis of the amino acid sequences confirmed that the 5 reprolysin-like and 3 astacin-like sequences contained their respective family's conserved zinc-binding His-motif as well as the conserved methionine, which clearly classified all eight *R. microplus* transcripts as metzincins. Phylogenetic analysis supports the alignment data and indicated that the 5 reprolysin-like and 3 astacin-like *R. microplus* MPs do belong to the reprolysin and astacin metzincin families, respectively.

Gene specific primers, which were specifically designed for each *R. microplus* metzincin MP, successfully amplified each transcript. Therefore it is evident that all 5 reprolysin-like transcripts and all 3 astacin-like transcripts are present in the RNA pool of *R. microplus* mixed life stages. The exact tissue and feeding life stage in which each transcript is expressed was investigated using RT-PCR and gel electrophoresis. Table 2.6 summarises the results of the expression profiling.Overall the results indicated that the reprolysins were most abundantly expressed in the salivary glands, whereas the astacins were most abundant in the midgut.

Table 2.6 A summary of the expression profiles of the 5 reprolysin-like and 3 astacin-like *R. microplus* MPs. Bold letters indicate the tissue in which a transcript is most abundantly expressed.

	Tissue	Life stage
BmMP1	Ov, **Sg**	Larvae, nymphs and adults
BmMP2	**Sg**	Larvae, nymphs and adults
BmMP3	Fb	Nymphs and adults
BmMP4	-	Larvae and nymphs
BmMP5	**Sg**	Larvae and female adults
As51	Acg, Sg, Mt, Ov, testis, **Mg**	All life stages
As70	Sg	Female adult
AsContig	Acg, Sg, Mt, Ov, testis, **Mg**	Larvae, nymphs and adults

The theoretical signal peptide, lack of GPI-anchor and transmembrane α-helix predictions suggest all 8 *R. microplus* metzincins to be secreted proteins. Therefore, by integrating all data it can be postulated that the 5 reprolysin-like MPs are secreted by the salivary glands into the saliva. Via the saliva these MPs are released into the

feeding cavity, where they most probably act on ECM components, to assist in maintaining a fluid intact feeding pool. For the astacin-like MPs it can be postulated that once a blood meal is obtained, the specialised secretory cells in the midgut epithelium release these enzymes. In the midgut lumen these astacin-like MPs most probably then act on ECM components to assure the blood does not clot within the midgut prior to digestion. However, the exact functions and substrates of these *R. microplus* metzincin MPs remain to be determined using precise activity assays of isolated or recombinant active proteins. However, to date the expression of recombinant and isolation of native MPs has been very limited.

What is of great interest is that this is the first report of metzincin MPs in ixodid tick tissues other than the salivary glands. Furthermore, this work presents, to our knowledge, the first description of astacin metalloproteases in ticks. Based on the data of this study and other invertebrate studies, it is evident that tick metzincins are not necessarily only involved in ECM digestion but may also be involved in other processes such as oogenesis, spermatogenesis and egg hatching.

It has also been speculated that metzincin metalloproteases may play a role in pathogen transmission (Francischetti *et al.*, 2003). It has in fact been shown that *Borrelia* spirochetes up-regulate the release and activation of matrix metalloprotease gelatinase B (MMP-9) and collagenase 1 (MMP-1) in human cells, enhancing the penetration of *B. burgdorferi* across ECMs (Gebbia *et al.*, 2001). Prevention of fibrin clot formation by salivary metalloproteases and other salivary anti-haemostatic agents (Ribeiro, 1989; Maritz-Olivier *et al.*, 2007) and dissolution of the fibrin clots eventually formed around the feeding cavity of ixodid ticks may similarly aid spirochete dissemination through vertebrate tissues.

2.5 References

Barrett, A. J., N. D. Rawlings and J. F. Woessner (2004). *Handbook of Proteolytic Enzymes*. Amsterdam, the Netherlands, Elsevier Academic Press.

Bjarnason, J. B. and J. W. Fox (1995). "Snake venom metalloendopeptidases; reprolysins." *Methods Enzymol* **248**: 345-368.

Blobel, C. P. (2005). "ADAMs: Key components in egfr signalling and development." *Nature Reviews Molecular Cell Biology* **6**(1): 32-43.

Bode, W., F.-X. Gomis-Rüth, R. Huber, R. Zwilling and W. Stöcker (1992). "Structure of astacin and implications for activation of astacins and zinc-ligation of collagenases." *Nature* **358**(6382): 164-167.

Bond, J. S. and R. J. Beynon (1995). "The astacin family of metalloendopeptidases." *Protein Science* **4**(7): 1247-1261.

Bowles, V. M., A. R. Young and S. C. Barker (2008). "Mettalloproteases and egg-hatching in *Pediculus humanus*, the body (clothes) louse of humans (Phthiraptera: Insecta)." *Parasitology* **135**(1): 125-130.

Brew, K. and H. Nagase (2010). "The tissue inhibitors of metalloproteinases (TIMPs): An ancient family with structural and functional diversity." *Biochim Biophys-Acta* **1803**: 55–71.

Chassy, B. M., A. Mercenier and J. Flickinger (1988). "Transformation of bacteria by electroporation." *Trends in Biotechnology* **6**(12): 303-309.

Decrem, Y., J. Beaufays, V. Blasioli, K. Lahaye, M. Brossard, L. Vanhamme and E. Godfroid (2008). "A family of putative metalloproteases in the salivary glands of the tick *Ixodes ricinus*." *Federation of European Biochemical Societies Journal* **275**(7): 1485-1499.

Decrem, Y., M. Mariller, K. Lahaye, V. Blasioli, J. Beaufays, K. Z. Boudjeltia, M. Vanhaeverbeek, M. Cérutti, L. Vanhamme and E. Godfroid (2007). "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**(5): 549-560.

Edwards, D. R., M. M. Handsley and C. J. Pennington (2009). "The ADAM metalloproteinases." *Molecular Aspects of Medicine* **29**(5): 258-289.

Francischetti, I. M. B., T. N. Mather and J. M. C. Ribeiro (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**(4): 869-875.

Gebbia, J. A., J. L. Coleman and J. L. Benach (2001). "*Borrelia* spirochetes upregulates release and activation of matrix metalloproteinase gelatinase B (MMP-9) and collagenase 1 (MMP-1) in human cells." *Infection and Immunity* **69**(1): 456-462.

Gerard, G. F., J. M. D'Alessio, M. L. Kotewicz and M. C. Noon (1986). "Influence on stability in *Escherichia coli* of the carboxy-terminal structure of cloned Maloney murine leukemia virus reverse transcriptase." *DNA* **5**(4): 271-279.

Gomis-Rüth, F.-X. (2003). "Structural aspects of the metzincin clan of metalloendopeptidases." *Molecular Biotechnology* **24**(2): 157-202.

Gomis-Rüth, F.-X. (2009). "Catalytic domain architecture of metzincin metalloproteases." *Journal of Biological Chemistry* **284**(23): 15353-15357.

Gomis-Rüth, F.-X., L. F. Kress and W. Bode (1993a). "First structure of a snake venom metalloproteinase: A prototype for matrix metalloproteinases/collagenases." *EMBO Journal* **12**(11): 4151-4157.

Hall, T. A. (1999). "BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." *Nucleic Acids Symposium Series* **41**: 95-98.

Harnnoi, T., T. Sakaguchi, Y. Nishikawa, X. Xuan and K. Fujisaki (2007). "Molecular characterization and comparative study of 6 salivary gland metalloproteases from the tick, *Haemaphysalis longicornis*." *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology* **147**(1): 93-101.

Hati, R., P. Mitra, S. Sarker and K. K. Bhattacharyya (1999). "Snake venom hemorrhagins." *Critical Reviews in Toxicology* **29**(1): 1-19.

Hite, L. A., J. D. Shannon, J. B. Bjamason and J. W. Fox (1992). "Sequence of a cDNA clone encoding the zinc metalloproteinase hemorrhagic toxin e from *Crotalus atrox*: Evidence for signal, zymogen, and disintegrin-like structures." *Biochemistry* **31**(27): 6203-6211.

Hooper, N. M. (1996). *Zinc metalloproteases in health and disease*, CRC Press.

Iba, K., R. Albrechtsen, B. J. Gilpin, F. Loechel and U. M. Wewer (1999). "Cysteine-rich domain of human ADAM 12 (meltrin α) supports tumor cell adhesion." *American Journal of Pathology* **154**(5): 1489-1501.

Jia, L.-G., K.-I. Shimokawa, J. B. Bjarnason and J. W. Fox (1996). "Snake venom metalloproteinases: Structure, function and relationship to the ADAMS family of proteins." *Toxicon* **34**(11-12): 1269-1276.

Kotewicz, M. L., J. M. D'Alessio and K. M. Driftmier (1985). "Cloning and overexpression of Moloney murine leukemia virus reverse transcriptase in *Escherichia coli*." *Gene* **35**(3): 249-258.

Lu, X., M. F. Scully and V. V. Kakkar (2005). "Snake venom metalloproteinase containing a disintegrin-like domain, its structure-activity relationships at interacting with integrins." *Current Medicinal Chemistry: Cardiovascular and Hematological Agents* **3**(3): 249-260.

Maritz-Olivier, C., C. Stutzer, F. Jongejan, A. W. H. Neitz and A. R. M. Gaspar (2007). "Tick anti-hemostatics: Targets for future vaccines and therapeutics." *Trends in Parasitology* **23**(9): 397-407.

Park, H. I. and L.-J. Ming (1998). "The mechanistic role of the coordinated tyrosine in astacin." *Journal of Inorganic Biochemistry* **72**(1-2): 57-62.

Porter, S., I. M. Clark, L. Kevorkian and D. R. Edwards (2005). "The ADAMTS metalloproteinases." *Biochemical Journal* **386**(1): 15-27.

Rawlings, N. D., F. R. Morton, C. Y. Kok, J. Kong and A. J. Barrett (2008). "MEROPS: The peptidase database." *Nucleic Acids Research* **36**: D320-D325.

Reiss, K. and P. Saftig (2009). "The "A Disintegrin And Metalloprotease" (ADAM) family of sheddases: Physiological and cellular functions." *Seminars in Cell and Developmental Biology* **20**(2): 126-137.

Ribeiro, J. M. C. (1989). "Role of saliva in tick/host interactions." *Experimental and Applied Acarology* **7**(1): 15-20.

Rychlik, W. and R. E. Rhoads (1989). "A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and *in vitro* amplification of DNA." *Nucleic Acids Research* **17**(21): 8543-8551.

Rychlik, W., W. J. Spencer and R. E. Rhoads (1990). "Optimization of the annealing temperature for DNA amplification *in vitro*." *Nucleic Acids Research* **18**(21): 6409-6412.

Sambrook, J., E. F. Fritsch and T. Maniatis (1989). *Molecular cloning: A laboratory manual*. New York, Cold Spring Harbor Laboratory Press.

Seals, D. F. and S. A. Courtneidge (2003). "The ADAMs family of metalloproteases: Multidomain proteins with multiple functions." *Genes and Development* **17**(1): 7-30.

Shimell, M. J., E. L. Ferguson, S. R. Childs and M. B. O'Connor (1991). "The Drosophila dorsal-ventral patterning gene *tolloid* is related to human bone morphogenetic protein 1." *Cell* **67**(3): 469-481.

Springman, E. B., E. L. Angleton, H. Birkedal-Hansen and H. E. van Wart (1990). "Multiple modes of activation of latent human fibroblast collagenase: Evidence for the role of a Cys73 active-site zinc complex in latency and a 'cysteine switch' mechanism for activation." *Proceedings of the Natural Academy of Sciences of the United States of America* **87**(1): 364-368.

Sterchi, E. E., W. Stöcker and J. S. Bond (2009). "Meprins, membrane-bound and secreted astacin metalloproteinases." *Molecular Aspects of Medicine* **29**(5): 309-328.

Stöcker, W., F.-X. Gomis-Rüth, W. Bode and R. Zwilling (1993). "Implications of the three-dimensional structure of astacin for the structure and function of the astacin family of zinc-endopeptidases." *European Journal of Biochemistry* **214**(1): 215-231.

Stöcker, W., F. Grams, U. Baumann, P. Reinemer, F.-X. Gomis-Rüth, D. B. McKay and W. Bode (1995). "The metzincins - Topological and sequential relations between the astacins, adamalysins, serralysins and matrixins (collagenases) define a superfamily of zinc-peptidases." *Protein Science* **4**(5): 823-840.

Stöcker, W., M. Ng and D. S. Auld (1990). "Fluorescent oligopeptide substrates for kinetic characterization of the specificity of *Astacus* protease." *Biochemistry* **29**(45): 10418-10425.

Takahara, K., G. E. Lyons and D. S. Greenspan (1994). "Bone morphogenetic protein-1 and a mammalian tolloid homologue (mTld) are encoded by alternatively spliced transcripts which are differentially expressed in some tissues." *Journal of Biological Chemistry* **269**(51): 32572-32578.

Tamura, K., J. Dudley, M. Nei and S. Kumar (2007). "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0." *Molecular Biology and Evolution* **24**(8): 1596-1599.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins (1997). "The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools." *Nucleic Acids Research* **25**(24): 4876-4882.

Tsai, P.-L., C.-H. Chen, C.-J. Huang, C.-M. Chou and G.-D. Chang (2004). "Purification and cloning of an endogenous protein inhibitor of Carp Nephrosin, an astacin metalloproteinase." *Journal of Biological Chemistry* **279**(12): 11146-11155.

Valenzuela, J. G., I. M. B. Francischetti, V. M. Pham, M. K. Garfield, T. N. Mather and J. M. C. Ribeiro (2002). "Exploring the sialome of the tick *Ixodes scapularis*." *Journal of Experimental Biology* **205**(18): 2843-2864.

Williamson, A. L., S. Lustigman, Y. Oksov, V. Deumic, J. Plieskatt, S. Mendez, B. Zhan, M. E. Bottazzi, P. J. Hotez and A. Loukas (2006). "*Ancylostoma caninum* MTP-1, an astacin-like mettaloprotease secreted by infective hookworm larvae, is involved in tissue migration." *Infection and Immunity* **74**(2): 961-967.

Wolfsberg, T. G., P. Primakoff, D. G. Myles and J. M. White (1995). "ADAM, a novel family of membrane proteins containing A Disintegrin And Metalloprotease domain: Multipotential functions in cell-cell and cell-matrix interactions." *Journal of Cell Biology* **131**(2): 275-278.

Young, A. R., N. Mancuso, E. N. T. Meeusen and V. M. Bowles (2000). "Characterisation of proteases involved in egg hatching of the sheep blowfly, *Lucilia cuprina*." *International Journal for Parasitology* **30**(8): 925-932.

Chapter 3

In vivo gene silencing of *Rhipicephalus microplus* metzincins

3.1 Introduction

In order to formulate an improved tick control strategy, it is of great importance to gain a comprehensive understanding of tick protein function and the significance of their roles in tick development, feeding and reproduction. Currently, one of the major limiting factors in tick research remains the lack of available genomic and proteomic information (de la Fuente and Kocan, 2006). In *R. microplus* the great genome size, of approximately 7.1 Gbp, and the high ratio of repetitive and exonic sequences greatly delays generation of the complete genome sequence (Moolhuijzen *et al.*, 2011). As a result the major genome resources, for this tick species, up to very recent remained limited to EST datasets. The *R. microplus* EST sequences have been obtained with the use of high through-put screening and various sequencing techniques (Wang *et al.*, 2007) and with the utilisation of molecular and bioinformatics tools several subsequent EST studies have provided information on the sequences of several genes encoding proteins of significant function. However, a great number of genes (64% of the BmiGI ESTs) could not be assigned gene ontology annotation and remained unclassified (Wang *et al.*, 2007). The recent release of the first draft results from the genome sequencing project of *R. microplus* (Bellgard et al., 2012) therefore provide a substantial platform for enhanced vaccine candidates identification and validation.

Evaluation methods such as Expression Library Immunisation (ELI) and protein isolation followed by direct vaccination trials are very effective methods, which can provide profound knowledge on possible tick protective antigens (de la Fuente *et al.*, 2005). These methods do, however, require tick infestation of a large number of animals resulting in experiments that are laborious, expensive and difficult to standardise, mostly due to animal-to-animal variance. RNA interference (RNAi) in ticks offers an attractive alternative experimental tool, which can be employed *in vitro*, *ex vivo,* or *in vivo,* to facilitate the investigation of the function of a gene, by observing the phenotypical effect of silencing (de la Fuente *et al.*, 2005; de la Fuente *et al.*, 2007b).

This powerful reverse genetic tool enables the screening of a large number of genes in a relatively short time and with minimal animal usage. The power and utility of RNAi, specifically for silencing the expression of any gene for which the sequence data is available, has driven its incredible rapid adoption in becoming one of the most widely used techniques in ticks (Ramakrishnan *et al.*, 2005; de la Fuente and Kocan, 2006). To date RNAi has been successfully used for the study of tick gene function in feeding (Karim *et al.*, 2004; Ramakrishnan *et al.*, 2005; de la Fuente *et al.*, 2006a; Decrem *et al.*, 2007; Nijhof *et al.*, 2007; Gao *et al.*, 2011) and reproduction (de la Fuente *et al.*, 2006a; Zhou *et al.*, 2006b; Nijhof *et al.*, 2007), as well as for the characterisation of genes involved in the tick-pathogen interface (Pal *et al.*, 2004; Ramamoorthi *et al.*, 2005; de la Fuente *et al.*, 2006b). RNAi has also been proven to be a fast effective tool for preliminary characterisation of tick protective antigens as well as antigen combinations (de la Fuente *et al.*, 2005; de la Fuente *et al.*, 2006c). This conserved post-transcriptional gene silencing method can greatly facilitate progress towards the development of anti-tick vaccines.

3.1.1 RNA interference – an overview

Broadly defined, RNAi comprises a class of processes that use short RNAs (21-25 nucleotides in length) to inhibit gene expression in a sequence specific manner by either inducing mRNA degradation or by physically blocking mRNA translation.

This gene regulatory phenomenon was first observed in petunias, when Napoli *et al.* (1990) discovered that introduction of a pigment-producing gene, under the control of a promoter, suppressed expression of both the introduced gene and the homologous endogenous gene, so-called co-suppression. It was only eight years later that Fire *et al.* (1998) discovered that injection of double stranded RNA (dsRNA) – a mixture of both sense and anti-sense strands of target mRNA, rather than either strand alone – into the gonads of the nematode *Caenorhabditis elegans* results in extremely potent silencing. With this revelation the door was opened for the discovery and development of RNAi as a reverse genetic tool. During their study of transgene and virus induced post-transcriptional gene silencing in tomato plants, plant virologists Hamilton and Baulcombe (1999) provided the first evidence that small RNA molecules are the active species of the RNAi process. Several subsequent *in vitro* and *in vivo* studies collectively

demonstrated and provided proof that the dsRNA is converted by an endonucleolytic mechanism, into small RNA intracellular signaling molecules (Tuschl *et al.*, 1999; Hammond *et al.*, 2000; Zamore *et al.*, 2000). The discovery of these RNAs, later termed short interfering RNAs (siRNAs), provided the first evidence for a universal biochemical pathway of the silencing phenomena in plants and animals. To date RNAi related events have been described in almost all eukaryotic organisms including plants, protozoa, fungi, nematodes, insects, arthropods and mammals (Agrawal *et al.*, 2003; de la Fuente *et al.*, 2007b; Bellés, 2010).

The best-studied natural function of RNAi is the role of RNAi in the immune system as a direct defense mechanism against viruses. From several recent studies it is clear that RNAi, directed against viruses, is an important component of the innate antiviral immunity in plants, fungi, arthropods and nematodes, however, it remains an open question as to whether this pathway might be a component of antiviral immunity in vertebrates (Cullen, 2006; Ding and Voinnet, 2007; Haasnoot *et al.*, 2007). Another natural function of RNAi seems to be the protection of the host's genome against invasive mobile genetic elements such as transposons (Sijen and Plasterk, 2003; Aravin *et al.*, 2007; Chung *et al.*, 2008). Intriguing recent results also indicate that RNAi-related pathways might act in the control of epigenetic modifications and regulations of heterochromatin (Matranga and Zamore, 2007; Obbard *et al.*, 2009). All of the latter remain to be investigated in ticks.

3.1.2 RNAi mechanism and machinery

Since the discovery of RNAi, numerous *in vivo* and *in vitro* studies have aimed at understanding the different pieces of the RNAi machinery. From a combination of results, a widely conserved three step mechanistic model for RNAi has been derived (Hammond, 2005; Liu and Paroo, 2010). The first step is referred to as the RNAi initiating step. It involves the binding of a RNA nuclease to a large dsRNA fragment and its cleavage into smaller double stranded siRNAs. The second step involves the loading of the siRNAs into a multi-nuclease complex and the final step is the subsequent degradation or blocking of a homologous mRNA (Figure 3.1).

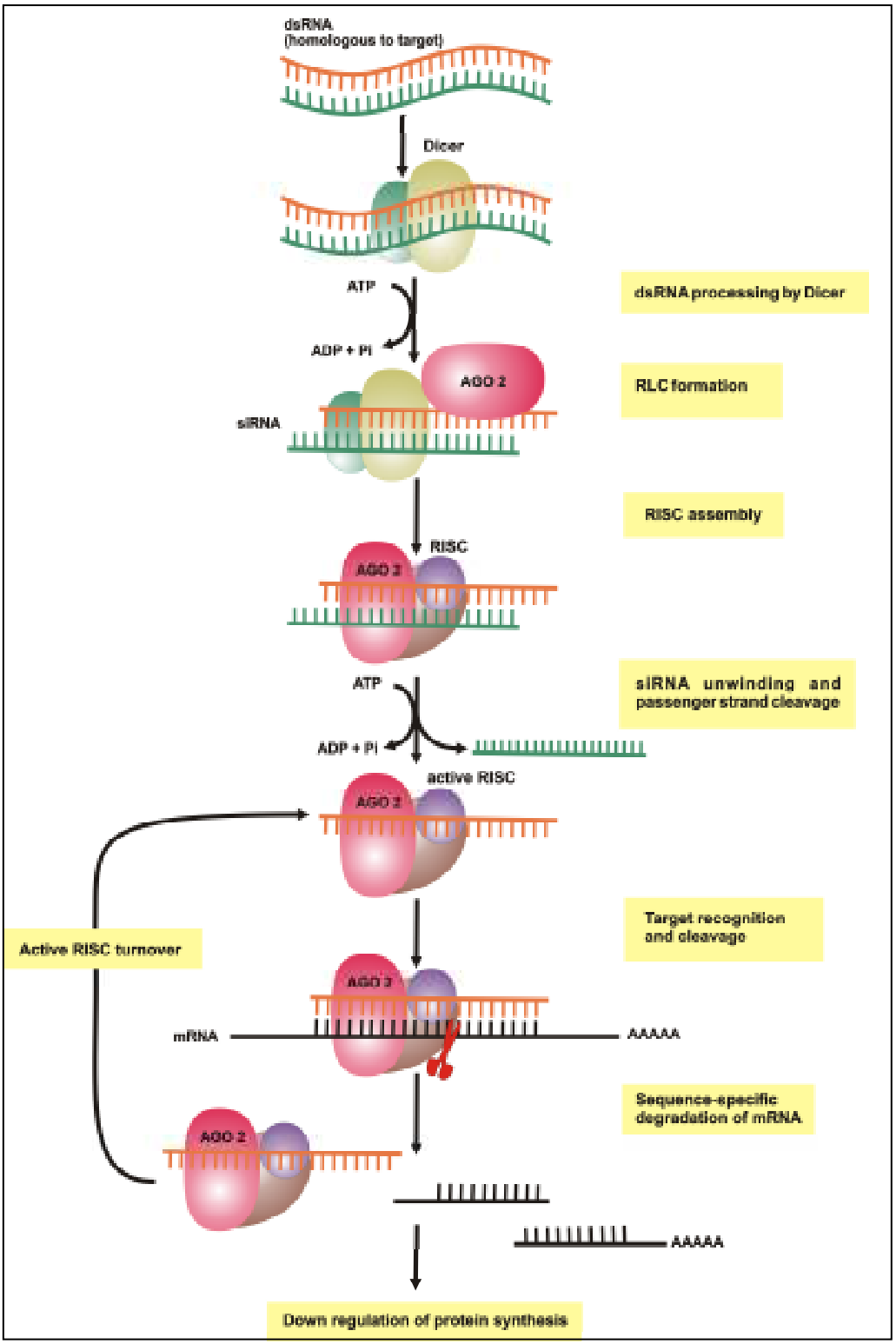

Figure 3.1. The RNAi process and the biochemical machinery involved (Tomari and Zamore, 2005; Sashital and Doudna, 2010). dsRNA is processed into short pieces (siRNA) by the endonuclease Dicer. The siRNA is loaded into the RNA-induced silencing complex (RISC), via the RISC loading complex (RLC), and the passenger strand is cleaved and released. The guide strand associates with a homologous mRNA strand by conventional base paring, and the mRNA strand is cleaved by RISC and released for further degradation within the cytoplasm.

3.1.2.1 Initiator step

The goal of the initiator step is the generation of siRNAs from dsRNA. When exploiting RNAi as a reverse genetic tool, the process is artificially induced by delivery of a dsRNA trigger. In non-mammalian systems long dsRNA (>200 bp) homologous to the target gene effectively triggers RNAi. In mammals, however, the introduction of dsRNA longer than 30 bp results in the activation of an anti-viral interferon response, which causes systematic non-specific inhibition of translation. Therefore in mammalian systems short synthetic siRNAs or DNA constructs, which express short hairpin RNAs (shRNAs), are typically used (McManus and Sharp, 2002; Agrawal *et al.*, 2003). Since the specificity of RNAi depends on the sequence and structure of the siRNAs (Tomari and Zamore, 2005), it is of great importance to carefully design the dsRNA or shRNA-constructs in order to maximise silencing of the target gene and ensure minimal off-target effects.

Once long dsRNA is introduced into non-mammalian cells it is processed into 21-25 nt double stranded RNAs (siRNAs), which have characteristic structures reflecting production by a ribonuclease III (RNaseIII) (Zamore *et al.*, 2000; Elbashir *et al.*, 2001; Hammond, 2005). Bernstein and colleagues identified and demonstrated that the enzyme responsible for the initiation of RNAi in *Drosophila* is in fact a RNaseIII-like protein and owing to its biochemical function they named the enzyme, Dicer (Bernstein *et al.*, 2001). To date, Dicer homologues have been identified in a wide variety of eukaryotes including, Excavata, Unikonta, Alveolata, Stramenopila and Archaeplastida (Knight and Bass, 2001; Golden *et al.*, 2002; Graham *et al.*, 2010).

A simplified model for the mechanism of Dicer is given in Figure 3.2. This model suggests that Dicer acts as a monomer: the two ribonuclease III (RNaseIII) domains forming a single processing center by associating in an intramolecular pseudo-dimer. Each domain hydrolyses a single RNA strand of the duplex to generate a new terminus. Due to the dimer alignment configuration, products with 2 nt 3'-overhangs are created (Zhang *et al.*, 2004). Dicer preferentially processes dsRNA from the ends of the substrates, with the Piwi/Argonaute/Zwille (PAZ) domain recognizing the 3'-overhangs (Zhang *et al.*, 2002; Song *et al.*, 2003; Yan *et al.*, 2003; Ma *et al.*, 2004).

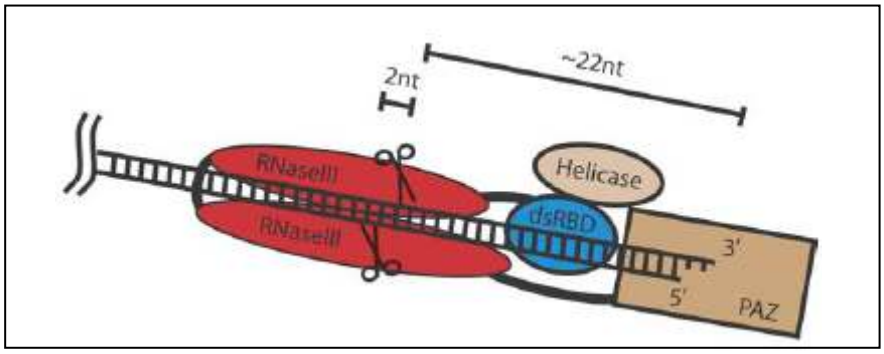

Figure 3.2. The proposed model for dsRNA catalysis by Dicer (Hammond, 2005). The PAZ domain binds the 2 nt 3'-overhang terminus of the dsRNA. The RNaseIII domains form a pseudo-dimer, with each domain hydrolysing one strand of the template. Two additional domains, a dsRNA Binding Domain (dsRBD) and putative helicase domain, are also presented in this model.

More profound insight into the structural mechanism used by Dicer came from the crystal structure of a Dicer protein derived from the protozoan *Giardia intestinalis* (MacRae *et al.*, 2006). This structure revealed that Dicer attains an elongated shape and by this means act as a molecular ruler, measuring the distance between the terminal binding PAZ domain and the active site, thereby giving rise to the 21-25 nt products (MacRae *et al.*, 2006).

Although the structure and mechanism of *Giardia* Dicer has been studied in detail, much less is known about Dicer proteins from higher eukaryotes (Lau *et al.*, 2009). Furthermore, *Giardia* Dicer is an atypical Dicer in that it is much smaller and simpler than any Dicer found in any other organism to date, containing only the two RNaseIII domains and a PAZ domain. For example, human Dicer (219 kDa) is nearly 3-times larger than *Giardia* Dicer (82 kDa). The difference in molecular mass is accounted for by at least five additional protein domains found in most Dicer proteins (Lau *et al.*, 2009). These include a TAR RNA-binding protein (TRBP) (Lee *et al.*, 2006), an amino terminal DExD helicase-like domain (Ma *et al.*, 2008), a putative dsRNA binding domain (dsRBD) named DUF283 (Dlakić, 2006), an Argonaute-binding domain (Sasaki and Shimizu, 2007) and a C-terminal dsRBD (Provost *et al.*, 2002). These additional domains participate in dsRNA processing, regulate dicing activity and serve as molecular scaffolds for consolidating protein factors involved in the initiation of RNAi (Lau *et al.*, 2009).

3.1.2.2 RNA-induced silencing complex (RISC) assembly

Once a dsRNA trigger has been processed into siRNAs, the guide strand (the strand which will provide sequence specificity for target mRNA cleavage) is identified and loaded into an effector ribonucleoprotein complex: the RNA-induced silencing complex (RISC). Since siRNAs cannot catalyse any reaction by themselves, RISC assembly is a key process for these small RNAs to exert their function (Kawamata and Tomari, 2010). To date, despite the importance of this process, it's exact mechanism remains limited. Based on several *in vitro* studies of RISC assembly, using *Drosophila* extracts and siRNAs as starting substrates, together with new structural studies of Argonaute (Ago), a revised 2-step model of the main features of this process has been derived (Jaskiewicz and Filipowicz, 2008; Wang *et al.*, 2009). The first step, known as RISC-loading, involves the insertion of the siRNA duplexes into the Ago-protein and the second step comprises the dissociation or unwinding of the passenger strand from RISC (Kawamata and Tomari, 2010).

In the model it is postulated that Dicer and a dsRNA binding partner (e.g. in *Drosophila* it is Dicer-2 and R2D2) are bound to the double stranded siRNAs, to direct it to the RISC complex (Tomari *et al.*, 2004a). Current findings suggest that Dicer-2 and R2D2 form a heterodimer that binds the siRNA, with Dicer binding the thermodynamically less stable end of the siRNA and the dsRBD domain binding the more stable end (Tomari *et al.*, 2004b; Kawamata and Tomari, 2010). These thermodynamics and binding asymmetries appear to determine which strand is finally incorporated into RISC. The guide strand is always the strand whose 5' end is less tightly paired to its complement, implying the strand with the least thermodynamically stable 5'end (Khvorova *et al.*, 2003; Schwarz *et al.*, 2003). Once the Dicer/dsRBD/siRNA ternary complex is formed it assembles with Ago-protein to form a RISC loading complex (RLC), which is an intermediate from which Dicer/dsRBD is gradually displaced by the Ago-protein to form pre-RISC (Pham *et al.*, 2004; Tomari *et al.*, 2004a; Kawamata and Tomari, 2010). Once bound to the siRNA duplex, Ago cleaves the passenger strand, triggering its dissociation (Rand *et al.*, 2005). After discharge of the passenger strand, possibly in an ATP-dependent step, the guide strand is in close association with Ago, forming a fully mature RISC, also referred to as siRISC (Tomari *et al.*, 2004b; Matranga *et al.*, 2005; Kawamata and Tomari, 2010).

3.1.2.3 Slicing/Silencing step

With chromatographic purification of RISC nuclease activity from *Drosophila* cells, several RISC components have been identified, however, only a few have been characterised at the functional level (van den Berg *et al.*, 2008). From current findings it is evident that RISC consists of several proteins and RNA molecules that all together act as the key actors of RNAi promoting mRNA degradation, repression of translation and remodeling of chromatin structure (Hutvagner and Simard, 2008). The central catalytic component of RISC, Argonaute 2, was the first identified constituent (Hammond *et al.*, 2001). Argonaute proteins are characterised by two conserved domains: the PAZ (also present in Dicer) and PIWI domains, respectively (Song *et al.*, 2004). Based on a higher degree of homology to either *Arabidopsis* AGO-1 or to *Drosophila* PIWI this family of proteins is subdivided into two highly conserved subclasses (Sasaki *et al.*, 2003). Recently an additional subclass, *C. elegans*-specific group 3 Argonautes, has been identified (Yigit *et al.*, 2006). Other components which have been identified as part of the RISC-complex include the Vasa intronic gene (VIG)-RNA binding protein, helicase proteins, Tudor- staphylococcal nuclease (Tudor-SN) and Fragile X -protein (dFXR) homologues (Caudy *et al.*, 2002; Caudy *et al.*, 2003), however, to date most of these proteins remain functionally uncharacterised. Therefore, due to the lack of structural data, the exact catalytic mechanism of target mRNA cleavage remains to be defined. A simplified model is given in Figure 3.3.

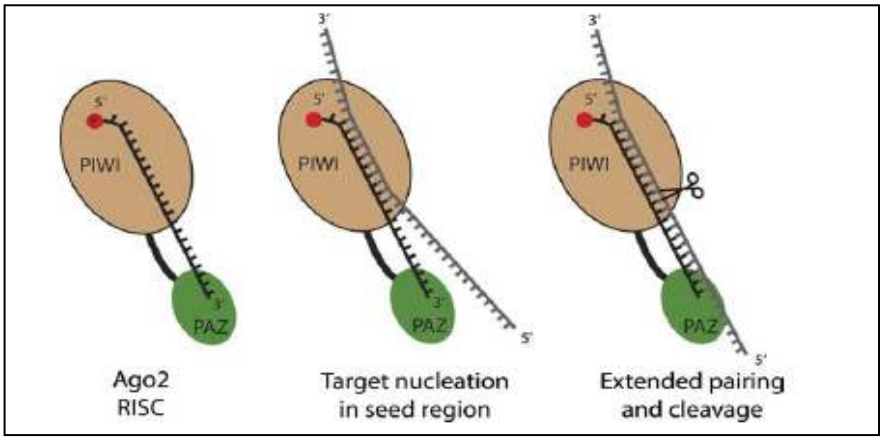

Figure 3.3. A model for targeted mRNA catalysis (Hammond, 2005). The guided strand is bound at the 3' end by the PAZ domain and at the 5' end by the PIWI domain. The targeted mRNA is initially bound by the seed region of the guide strand and then subsequent binding extends to the 3'end. The RNaseH fold within the PIWI domain hydrolyses the mRNA. The product is released and further degraded.

In the RISC complex the 3' end of the siRNA guide strand is bound by the PAZ domain, while the PIWI domain contacts the 5' phosphorylated end (Song *et al.*, 2003; Yan *et al.*, 2003; Lingel *et al.*, 2004; Ma *et al.*, 2004). Structural studies of prokaryotic Ago-like protein in complex with siRNA mimics demonstrated that the sugar phosphate back bone of the 5' end of the guide strand (nucleotide 2 to 8) contact the PIWI domain in such a manner that the corresponding bases are presented on the surface (Ma *et al.*, 2005). This region is known as the 'seed' region (Hammond, 2005). Corresponding mRNA targets are initially bound by the seed region of the siRNA and then pairing is extended to the 3' end. A single cleavage of the target mRNA occurs across from nucleotide 10 and 11 (with respect to the 5' end) of the guide stand. The catalytic engine is an RNaseH fold present in the PIWI domain (Parker *et al.*, 2004; Song *et al.*, 2004; Ma *et al.*, 2005). Although target cleavage by Argonaute does not require ATP *per se*, release of the cleaved target (RISC recycling) is stimulated by ATP (Nykänen *et al.*, 2001; Haley and Zamore, 2004). From *in vivo* studies it is evident that, once released from RISC, the 3' mRNA fragment is degraded in the cytoplasm by exonuclease *Xrn*1, while the 5' fragment is degraded by a complex of exonucleases, the so-called exosome (Zamore and Haley, 2005). In the RNAi pathway, siRISC operates as a multiple turn over enzyme. Once the mRNA target is cleaved, siRISC dissociates from the cleaved mRNA to repeat another cleavage cycle (Tomari and Zamore, 2005).

3.1.3 RNA interference – in ticks

Up to 2007 only one RNAi tick protein, a putative *R. microplus* Ago-2, had been identified (de la Fuente *et al.*, 2007b). Therefore, in order to gain a better understanding of the tick RNAi mechanism Kurscheid and colleagues set out in 2009 to identify more tick homologues of the RNAi machinery. With the use of several bio-informatic tools, they screened more than 13 000 *R. microplus* ESTs together with *I. scapularis* genome reads (from the first Chelicerata genome project). Thirty one putative tick RNAi proteins were identified including: a Dicer, RISC associated Ago-2 and FmRp, RNA dependent RNA polymerase (EGO-1) and 23 homologues implicated in dsRNA uptake and processing (Table 3.1.) (Kurscheid *et al.*, 2009).

Table 3.1. Putative tick RNAi candidate homologues (adapted from Kurscheid *et al.*, 2009)

Function	Protein*	*R. microplus* BmiGI (% identity)[a]	*I. scapularis* contig ID (% identity)[b]
DICER			
RNase III dsRNA processing	*Dcr-1 *Cele*	incomplete	DS643033 (31%)
ARGONAUTE proteins – target recognition and silencing			
RISC – miRNA pathway	Ago-1	TC13769 (43% DUF1785 and PAZ domains),	ABJB010009424.1 (41%)
		TC6448 (44% PIWI domain)	
RISC – RNAi pathway	Ago-2	TC8091 (25% DUF1785 & PAZ domains)	ABJB010128003.1 (31%)
Systemic RNAi			
Proposed role in transport of dsRNA	*Rsd-3 *Cele*	MPAAN09TR (45%)	ABJB010279725.1 (48%)
dsRNA uptake and processing			
Endocytic protein (EPsiN)	*Epn-1*Cele*	BEADR88TR (71%)	ABJB010748067.1 (43%)
Vesicle mediated transport	AP-50	C6127 (89%)	TABJB010508398 (91%)
	Arf72	Not found	ABJB010115816.1 (68%)
	Clathrin hc	TC10346 (60%)	ABJB010065986.1 (87%)
Endosome transport	Rab7	BEAGW52TR (80%)	ABJB010159881.1 86%)
Intracellular transport	CG3911	TC6954 (67%)	ABJB010384785 (64%)
	Cog3	TC5984 (49%)	ABJB010296208.1 (68%)
	IdlCp	Not found	ABJB011123114.1 (52%)
Lysosomal transport	Lt	TC12854 (35%)	Not found
Lipid metabolism	Gmer	TC9381 (62%)	Not found
	Pi3K59F	BEADT89TR (52%)	Not found
	Sap-r	TC9046 (22%)	Not found
Proteolysis and peptidolysis	CG4572	TC6395 (35%)	ABJB010180836.1 (43%)
	CG5053	Not found	ABJB010804049.1 (63%)
	CG8184	Not found	ABJB010385401.1 (72%)
Oogenesis	Egh	TC8075 (67%)	ABJB010259843.1 (66%)
Rhodopsin mediated signaling	ninaC	Not found	ABJB011087029.1 (42%)
Translation regulation	Srp72	Not found	ABJB010441811.1 (38%)
ATP synthase/ATPase	Vha16	MPAA174TR (59%)	ABJB010975295.1 (67%)
	VhaSFD	TC10823 (63%)	ABJB010753004 (56%)
Unknown	CG5161	TC14816 (61%)	Not found
	CG5382	Not found	ABJB010478954.1 (84%)
Other factors			
RISC assembly	Armitage	TC9347 (35%)	Not found
RISC associated nuclease	TudorSN	BEAFW62TR (46%)	ABJB010481234.1 (48%)
RISC function	FMRp	BEAE145TR (53%)	ABJB010028120.1 (67%)
ATP-dependent RNA helicase	Rm62	TC14966 (70%)	ABJB010043214.1 (54%)
RNA-directed RNA polymerase	*EGO-1*Cele*	BEAEL55TR (41%)	ABJB010057970.1 (54%)

*All proteins of the RNAi pathway originate from *D. melanogaster* except those indicated as *cele* (that are derived from *C. elegans).*

[a] Sequences for *R. microplus* were obtained from the *Rhipicephalus microplus* Gene Index (BmiGI) at http://combio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=bmicroplus/.

[b] *I. scapularis* genome project (IGP) data was accessed through http://www.ncbi.nlm.nih.gov/sites/entrez/

The recent comparative genomics studies between the processes of RNAi in *D. melanogaster* and the two tick species, *R. microplus* and *I. scapularis*, provided evidence that even though many of the basic RNAi machinery are evolutionary conserved, several components differ, indicating that the tick RNAi pathway may be different from that of other arthropods such as insects (Kurscheid *et al.*, 2009).

In various plant and nematode species it has been found that the silencing signal is amplified by a process reliant on a RNA-dependent RNA polymerase (RdRP), which in effect permits a systemic RNAi response from the site of origin to eventually most or all other tissues (Cerutti, 2003; Carthew and Sontheimer, 2009). A two-step mechanism has been described for RdRP amplification in *C. elegans* (Sijen, 2003). Once the primary siRNA has been generated by Dicer, RdRP uses the guide siRNA strand as primer and native mRNA as template to synthesize abundant secondary siRNAs and increasing the effectiveness of RNAi (Figure 3.4). As RNA synthesis occurs in a 5'–3' direction, this amplification leads to the 5' spreading of the initial RNA interference signal. This phenomenon is known as transitive RNAi (Sijen, 2003). To date this RdRP-dependent RNAi amplification seems absent in higher eukaryotes (including *D. melanogaster* and mosquitoes) (Bellés, 2010).However, based on several tick RNAi studies it is evident that this phenomena is present in ticks, since with the application of dsRNA via body cavity injection, feeding or soaking, global and persistent gene silencing in treated ticks and their progeny occurs (Karim *et al.*, 2004; Soares *et al.*, 2005; Nijhof *et al.*, 2007). Recent preliminary studies have indicated that this systemic RNAi effect in ticks is most likely attributed to (Kurscheid *et al.*, 2009) a putative homolog of the RdRP EGO-1 protein from *C. elegans* has also been identified in the tick species *I. scapularis*. This protein is associated with the amplification of trigger RNA by transcribing additional dsRNA molecules using target mRNA as a template, thereby yielding secondary siRNAs and increasing the effectiveness of RNAi (Saleh *et al.*, 2006; Aung *et al.*, 2011).

In order for RNAi to spread from one cell to another a transport system is required. The phenomenon of the progression of an RNAi response from the site of origin to neighboring cells and eventually to all, or most, tissues of the organism was first described in plants and *C. elegans*. The systemic response may even persist through multiple developmental stages, including being transferred through the germline to the progeny. It has been proposed that this response has evolved as a strategy to increase

the silencing efficiency (Buchon and Vaury, 2006) and it is known to form part of the immune system in plants. In *C. elegans* several proteins have been identified, which are necessary for the systemic RNAi spread, these include: a multi-transmembrane protein (SID-1: systemic RNA interference deficient-1) thought to act as a channel for dsRNA uptake (Winston *et al.*, 2002; Jose *et al.*, 2009; Calixto *et al.*, 2010); as well as RNAi spreading defective proteins (RSD2, RSD3, RSD6) (Tijsterman *et al.*, 2004). To date only RSD3 and Endocytic protein (Epn-1) have been identified in ticks, leaving the mechanism of cell to cell dsRNA transport in ticks unresolved.

An application made possible, is the study of genes in life stages such as eggs that may not be amenable to other methods of dsRNA delivery. By delivering the relevant dsRNA to a preceding life stage, in this case the adult female, systemic RNAi will propagate the necessary signal throughout the adult as well as within any eggs produced by that female (Nijhof *et al.*, 2007). This application is known as parental, trans-stadial or transovarial RNA interference. Exploiting systemic RNA interference during RNAi studies is also a relatively simple procedure.

3.1.4 Tick RNAi – mechanism and machinery

Although more pieces of the tick RNAi machinery have been identified, the exact mechanism remains unknown. The current model for tick RNAi is based on integrated knowledge of the RNAi process of other organisms including, *D. melanogaster, C. elegans* and mosquitoes (Figure 3.4.).

The postulated mechanism begins with the delivery of dsRNA (from either an exogenous or viral source) into the cytoplasm. It is hypothesised that this process is mediated by a protein similar to the *C. elegans* sid-1 protein (de la Fuente *et al.*, 2007b). However, a second pathway for dsRNA uptake has been described for other invertebrates (including *D. melanogaster*) namely, the 'alternative' endocytosis-mediated uptake mechanism (Saleh *et al.*, 2006; Huvenne and Smagghe, 2010). Once the dsRNA have entered the cytoplasm it is processed into double-stranded siRNAs, ~21-23 nt in length, by Dicer (as previously discussed in section 3.2.1.1). Although no *R. microplus* Dicer-homologue has been found to date, putative Dicer activity has been

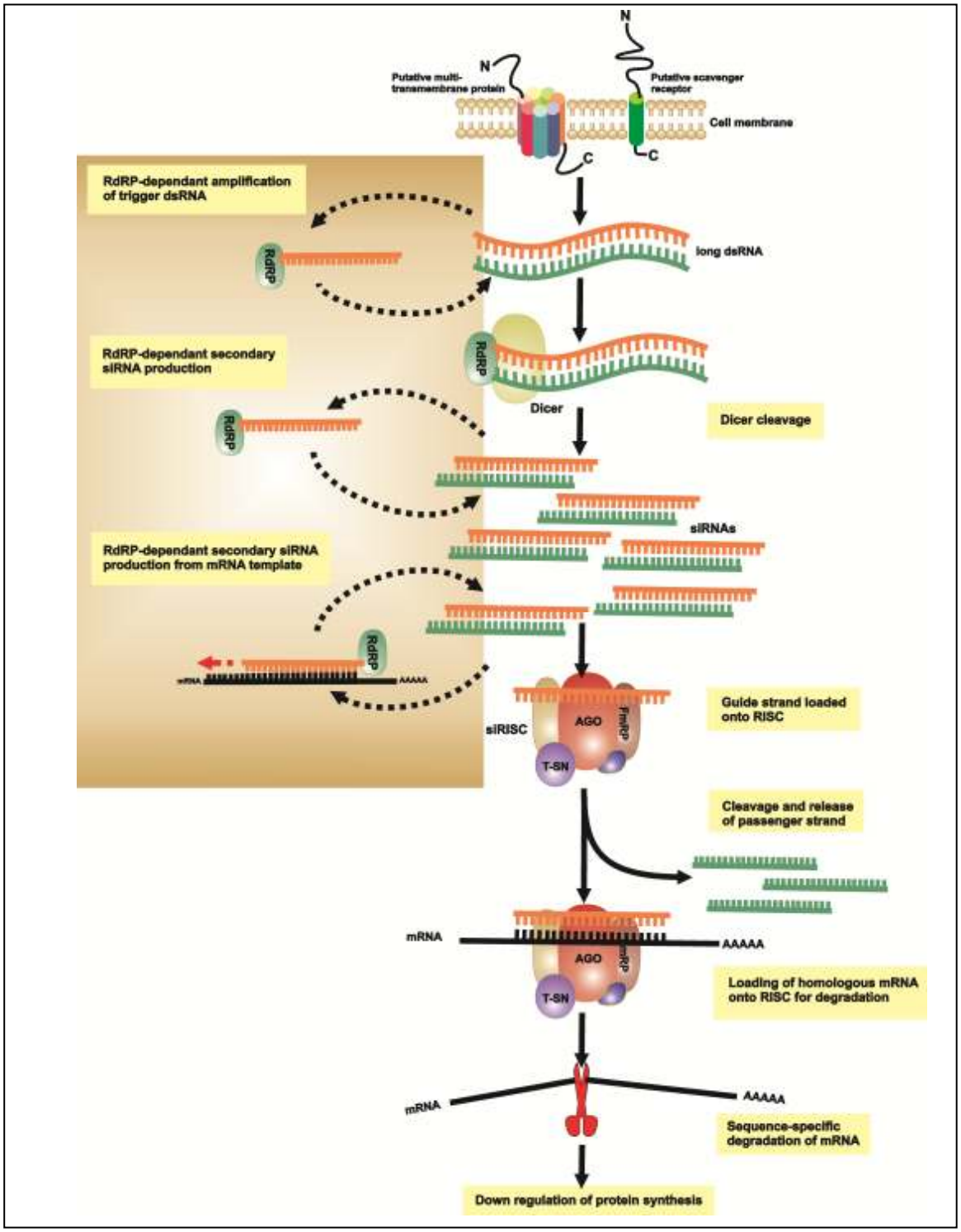

Figure 3.4. A schematic representation of a putative tick RNAi pathway (Barnard *et al.*, 2012). The proposed model might either use a multi trans-membrane protein (similar to SID-1) for dsRNA uptake or an endocytosis-mediated process which may include scavenger receptors. Once the dsRNA is in the cytoplasma it is processed into siRNAs ~ 21–23 nt in length, by a Dicer homologue. The siRNA are then presented to RISC which incorporates the siRNA, targets and degrades any homologous mRNA. RISC includes Ago, TudorSN and FmRp homologues. The proposed activity of RdRP is indicated as amplifying either trigger dsRNA, cleaved siRNA or using primary siRNA to prime synthesis of secondary siRNAs utilizing native mRNA as template. This causes 5' amplification and spreading of the initial RNAi signal, and is known as transitive RNAi.

observed when *R. microplus* protein extracts were incubated with dsRNA (personal communication with Dr. C. Maritz-Olivier). In addition, a Dicer homologue with expected domain structure for eukaryotic Dicer has been identified in *I. scapularis* in a recently assembled contig (GenBank: DS643033). From the current tick resources no definitive dsRBP, which usually associates with Dicer (such as *D. melanogaster* R2D2 or *C. elegans* Rde-4), could be identified (Kurscheid *et al.*, 2009).

After the initiation step, Dicer presents siRNAs to RISC, which then incorporates the siRNAs and subsequently targets and degrades any mRNA with a cognate sequence. To date several putative tick RISC components have been identified including: a *R. microplus* Ago-2 homologue (BmiGI: TC984), a tick TudorSN (*I. scapularis* tudor-staphylococcal nuclease, GenBank EEC18716.1) and a *R. microplus* FmRp (representing the *D. melanogaster* orthologue of the fragile-X mental-retardation protein essential to RISC). Based on their recent studies, Kurscheid *et al.* (2009) also suggested that within this model potential amplification of both trigger dsRNA and/or secondary siRNAs might take place. This reaction is probably mediated by a putative RdRP homologue and together with associated proteins facilitates systemic RNAi spread to other tissues and subsequent tick stages.

Clearly the RNAi pathway warrants further elucidation. More profound knowledge of sequence information (such as tick specific genome projects) and functional studies will create improved and efficient resources for the identification and characterisation of components of the RNAi machinery in ticks. This will in effect be beneficial for tick research and will greatly enhance the development of improved tick control measures.

3.1.5 Methods for RNAi in ticks

There are four widely used methods to deliver dsRNA to ticks for RNAi namely: (i) injection, (ii) soaking/ incubation, (iii) feeding and (iv) the production of dsRNA by a virus (de la Fuente *et al.*, 2007b). Injection of dsRNA is the most common used method for *in vivo* RNAi tick studies and can either be done manually (using a Hamilton syringe with a one inch, 33-gauge needle) or by means of microinjection. When dsRNA is delivered by way of manual injection, it is administered to the lower quadrant of the ventral surface of the tick's exoskeleton. With microinjection the dsRNA is administered

to the ventral torso of the idiosoma, away from the anal opening, with the use of a microdispensor and a micromanipulator connected to a microinjector. Even though manual injection of dsRNA has been proven to be a reliable and reproducible method in various tick species, microinjection is a more controlled method due to its ability to better control the injection volume. Both manual- and microinjection seem to yield similar survival rates in injected ticks (de la Fuente *et al.*, 2007b).

Soaking or incubation of ticks in dsRNA solutions offers a method by which RNAi can be applied to live immature ticks, tick tissues and cells. This method has been successfully used in *ex vivo, in vivo* and *in vitro* experiments (Bowman and Sauer, 2004; de la Fuente *et al.*, 2007b). If standardised it can provide a means in which gene silencing is induced relatively fast with minimum effort, in large numbers of immature ticks. Soaking/ incubation provides the best approach for *ex vivo* studies of gene expression in isolated tick tissues and in addition offers a way in which RNAi silencing can be induced to tick cell lines for high throughput screening and functional genomics.

dsRNA can also be delivered to ticks via feeding (Kocan *et al.*, 2005), by placing a capillary tube directly over the hypostome. Although this method is labour intensive and time consuming, it can prove valuable in gaining insight into tick–pathogen and tick–host interfaces. dsRNA can also be introduced into tick cells by infecting the cells with vector-borne RNA viruses. Once the dsRNA is produced by the virus, the RNAi pathway is triggered within the cell. For example, with recombinant Semliki Forest virus expressing RNA sequences of the tick-borne Hazara virus (HAZV), Garcia *et al.* (2005) were able to inhibit replication of HAZV. Finally, Karim *et al.* (2010) have recently desribed a delivery method for dsRNA to eggs and nymphs of *I. scapularis* using electroporation.

3.1.6 Transovarial RNAi

Transovarial RNAi is a novel, effective method which allows gene-specific silencing in eggs and subsequent immature lifestages of ixodid ticks. This method was developed and first employed by Nijhof *et al.* (2007), showing that silencing of tick protective antigens Bm86, Bm91 and subolesin can be induced in the eggs and larvae of *R. microplus* by injecting dsRNA into repleted female ticks.

During transovarial silencing, dsRNA is injected through a spiracle plate of a replete female tick according to a method first described by Bechara *et al.* (1988). The spiracle plates (the trachea openings) are sclerotised structures localised posterior to the fourth pair of legs (Sonenshine, 1991). Their inelastic structure allow for puncturing and injection of small quantities of fluid by a fine needle, without subsequent reflux of the injected solution, haemolymph or tissue. It has been established that upon injection of dsRNA into the tick haemocoel with this method, the course of oviposition remains unaffected and insignificant female tick mortality is observed (Nijhof *et al.*, 2007).

With the first transovarial RNAi study, Nijhof *et al.* (2007) found that eggs oviposited by females injected with dsRNA subolesin displayed an anomalous phenotype: undifferentiated cell masses instead of differentiated embryos were observed. With quantitative real-time PCR, gene-specific silencing within the eggs was confirmed. Gene-specific silencing was also observed throughout the larva lifestage, however, the effect diminished over time and finally in the subsequent adults no silencing was detected. Transovarial RNAi has also been successfully applied in 3-host ixodid ticks (*Amblyomma americanum, Dermacentor variabilis* and *I. scapularis*) in which similar trends of gene-specific silencing were observed (Kocan *et al.*, 2007). It is therefore suggested that a common transovarial RNAi mechanism exists throughout ixodid ticks. With real-time PCR, using primers designed within the dsRNA region, it has been shown that the unprocessed dsRNA is incorporated into the eggs. Therefore, it seems as if the suggested route of incorporation of exogenously produced yolk directly from haemolymph into oocytes may be followed for the incorporation of dsRNA into oocytes, rather than the germ-line spread found in *C. elegans* (Saito *et al.*, 2005).

Although limited to the egg and larvae stage, this gene silencing mechanism provides a simple and effective method for the rapid characterisation of ixodid tick genes involved in oviposition, embryogenesis and larval development.

3.1.7 Hypothesis

Reprolysin and astacin metzincins play key roles in vital physiological processes such as feeding, digestion, ovipositioning and embryogenesis.

3.1.8 Aims

- *In vivo* gene silencing (RNAi) of the five reprolysin and three astacin transcripts in feeding *R. microplus* adult female ticks, in order to investigate the phenotypical outcome.
- Transovarial gene silencing of the five reprolysin and three astacin transcripts in fully engorged *R. microplus* females, to investigate the role of these metzincins in ovipositioning and embryogenesis.
- To determine the percentage silencing of each transcript brought about by RNAi, using semi quantitative real-time PCR.
- To investigate the differential transcriptional responses between these metzincin family members.
- To determine the subsequent effect of RNAi on metalloprotease activity in the investigated tissues, using a metalloprotease specific assay.

3.2 Materials and Methods

3.2.1 Experimental animals

All *in vivo* RNAi studies were done at the Utrecht Centre for Tick-borne Disease (UCTD). Two male Holstein–Friesian calves (#87766 and #03903), 5 months old, which had no previous exposure to ticks, were used for tick feeding. All tick feedings were approved by the Animal Experiments Committee (DEC) of the Faculty of Veterinary Medicine, Utrecht University (DEC No. 2008.II.07.068).

3.2.2 Ticks and tick feeding

R. microplus ticks, which were initially collected in Mozambique, were provided by ClinVet international (Pty), Bloemfontein, South Africa and were subsequently bred and maintained at the UCTD, University of Utrecht. For tick feeding, patches with inner dimensions of 60 x 85 mm - sewn to open cotton bags - were utilised (Nijhof *et al.*, 2007). Five patches were glued to the shaved back of calf #87766 and two patches to calf #03903. Two batches of larvae (obtained from 1500 mg of pooled eggs, oviposited by 25 females) were placed on 2 different patches on day 0 (patch 1) and day 1 (patch 2) on calf #87766. Approximately 400 unfed females were collected on day 16 and 200 unfed males were collected on day 16 and day 17, respectively. Collected ticks were incubated at 27ºC with 95% relative humidity. On day 16 the freshly molted females were injected with prepared dsRNA or injection buffer as described below.

3.2.3 dsRNA preparation

3.2.3.1 Synthesis of template DNA

The first step in order to synthesise dsRNA (which corresponded to the target sequences) was to prepare DNA templates with single T7 RNA polymerase promoter sequences at the 5’ ends of each strand. The T7 promoter sequences were

incorporated into primer sequences and subsequently into DNA using PCR. With this method the generated PCR product (T7 incorporated DNA) can be directly used in a single transcription reaction to synthesise dsRNA, with the T7 sequences serving as binding sites for the RNA polymerase (Figure 3.5.).

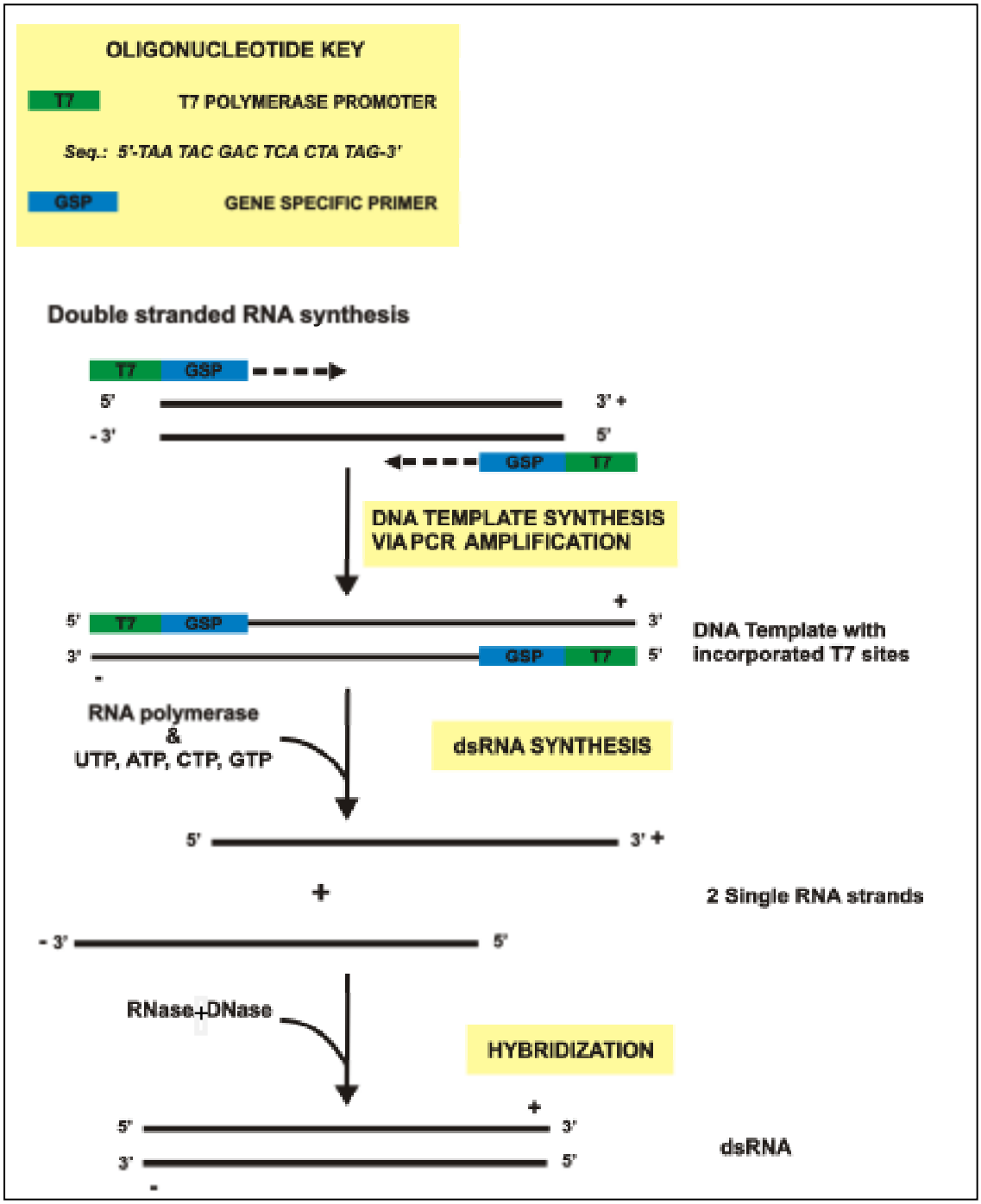

Figure 3.5. Double stranded RNA synthesis. The minimal sequence requirement for a T7 polymerase promoter (A). First DNA with incorporated 5' T7 promoter sites is synthesised. RNA polymerase can then recognise the T7 promoter sites and together with NTPs and hybridisation form dsRNA (B).

For each of the 5 reprolysin and 3 astacin *R. microplus* homologues a forward and reverse gene specific primer was designed, each with a T7 promoter sequence appended to the 5' end. The annealing temperature of each primer was calculated (as described in section 2.2.3) and the primers were synthesised by Invitrogen™ (The Netherlands). The primers were dissolved in double distilled deionised water and stock solutions, with final concentrations of 100 µM, were prepared.

For each *R. microplus* metzincin transcript three identical 50 µl PCR reactions were performed to generate sufficient dsDNA template for dsRNA synthesis. All PCR reactions were performed in 200 µl thin wall tubes and were conducted in an iCycler Thermal Cycler (Bio-Rad Labrotories, Inc., USA). Single stranded cDNA synthesised by Dr. A. Nijhof from *R. microplus* mixed lifestages or a 1/10 000 dilution of recombinant plasmid served as template. Two microlitre of template was added to a reaction solution containing final concentrations of 0.4 µM forward and reverse T7- gene specific primer, 1 mM Mg^{2+}, 0.1 mM of each dNTP, 1x Go*Taq* Buffer and 1.25 U of Go*Taq* (Promega, USA). The cycling reaction consisted of an initial denaturing step (94 °C, 2 minutes), followed by 40 cycles of denaturation (94°C, 30 sec), annealing (55°C, 30 sec), extension (72°C, 2 minutes), and a final extension step (72°C, 7 minutes). The PCR products were analysed on a 2% w/v agarose/TAE/EtBr gel (as described in section 2.2.6.).

3.2.3.2 Purification of PCR products (T7 incorporated DNA)

To purify the PCR products from unincorporated nucleotides, remaining polymerase, primers and salts, the pooled PCR reaction mixtures were subjected to column chromatography using the High Pure PCR Cleanup Micro Kit™ (Roche Diagnostics, Germany).

Purification of the PCR products was conducted with the High Pure PCR Cleanup Micro Kit™ according to the manufacturer's protocol (Roche Diagnostics, Germany). To each 100 µl pooled PCR reaction mixture, 400 µl Binding Buffer was added. The reaction mixture was vortexed and centrifuged briefly and the reaction mixture was transferred to a High Pure Filter Tube. The column was centrifuged in a collection tube for 60 sec at 8000 x g and the flow through was discarded. The column was then washed twice with

400 µl Wash Buffer with centrifugation at 8000 x g for 60 sec and 300 µl Wash Buffer with centrifugation at 8000 x g for 60 sec. The column was centrifuged an additional 1 minute at 16 000 x g in order to remove any excess alcohol. The column was placed in a clean 1.5 ml tube and 20 µl Elution Buffer was loaded onto the column and the column was incubated at 37°C with shaking for 5 minutes and subsequently centrifuged at 8000 x g for 1 minute. All PCR products' concentrations were determined with the NanoDrop ND-1000 spectrophotometer (Thermo-scientific, USA).

3.2.3.3 Simultaneous synthesis of both single stranded RNA (ssRNA) strands

To synthesise dsRNA from the T7 incorporated DNA template a single transcription reaction was conducted utilising the T7 Ribomax Express RNAi system (Promega, USA). This kit makes use of a T7 polymerase, which together with RNA nucleotides (ATP, GTP, CTP and UTP) synthesises two complementary RNA transcripts from the template (Figure 3.5.). Briefly, 1 µg of DNA template was added to 10 µl of RiboMax™ Express T7 2x Buffer, 2 µl Enzyme Mix, T7 Express and water to a final volume of 20 µl. The complete reaction mixture was incubated at 37°C for 16 hours.

3.2.3.4 Annealing of dsRNA and removal of DNA template and ssRNA

Following the transcription reaction the ssRNA strands were hybridised to form dsRNA. This step is dependent on a thermal gradient. Samples were incubated at 70°C for 10 minutes and gradually cooled down to room temperature, over a period of 20 minutes. Any remaining DNA and ssRNA were subsequently removed with nuclease digestion. One microlitre RQ1 RNase free DNase (1 U/ µl) and 1 µl of 1/200 dilution of an RNase A solution were added to each 20 µl reaction and the reactions were incubated at 37°C for 30 minutes.

3.2.3.5 Purification of synthesised dsRNA

The synthesised dsRNA was precipitated with sodium acetate. To each sample one tenth of the volume sodium acetate (3 M, pH 5.2) and 2.5 volumes of 95% ethanol were added. Reaction mixture was mixed, incubated on ice for 5 minutes and centrifuged for

10 minutes at 16 000 x g (4°C). The supernatant was discarded and the precipitate was washed with 500 µl cold 70% ethanol and subsequent centrifugation. Pellets were air dried and re-suspended in 40 µl nuclease-free water (Fermentas GmbH, Germany). A 1/50 dilution of each sample was analysed on a 2% w/v agarose/TAE/EtBr gel (as described in section 2.2.6.) and the final concentration of the synthesised dsRNA was determined with the NanoDrop ND-1000 spectrophotometer (Thermo-scientific, USA). The number of standard RNA molecules per microlitre was calculated with the following formula:

(Xg/µl RNA/ [transcript length in nucleotides x 640]) x 6.022 x 10^{23} = Y molecules/ µl

3.2.4 Injection of ticks with dsRNA

Five groups of 80 freshly molted females were injected, each group with a different combination of dsRNA (Table 3.2.).

Table 3.2. Injection combinations

Group	Silenced Transcripts
TE	Injection buffer alone
A	BmMP4, BmMP5
B	BmMP1, BmMP2, BmMP3
C	As51, As70
D	AsC

The female ticks were placed on double-sided tape with the ventral side upwards. After the lower right quadrant of the ventral surface of the exoskeleton was pierced with a 27G needle, 0.5 µl dsRNA solution (1-2 × 10^{12} molecules/ µl) was injected (Figure 3.6.). The dsRNA was dissolved in injection buffer (10 mM Tris-HCl, pH 7 and 1 mM EDTA) and the injection was carried out using a 10 µl syringe with a 33 G needle (Hamilton, Bonaduz, Switzerland) mounted on a MM3301-M3 micromanipulator (World Precision Instruments (WPI), Berlin, Germany) and connected to an UMPII syringe pump (WPI).

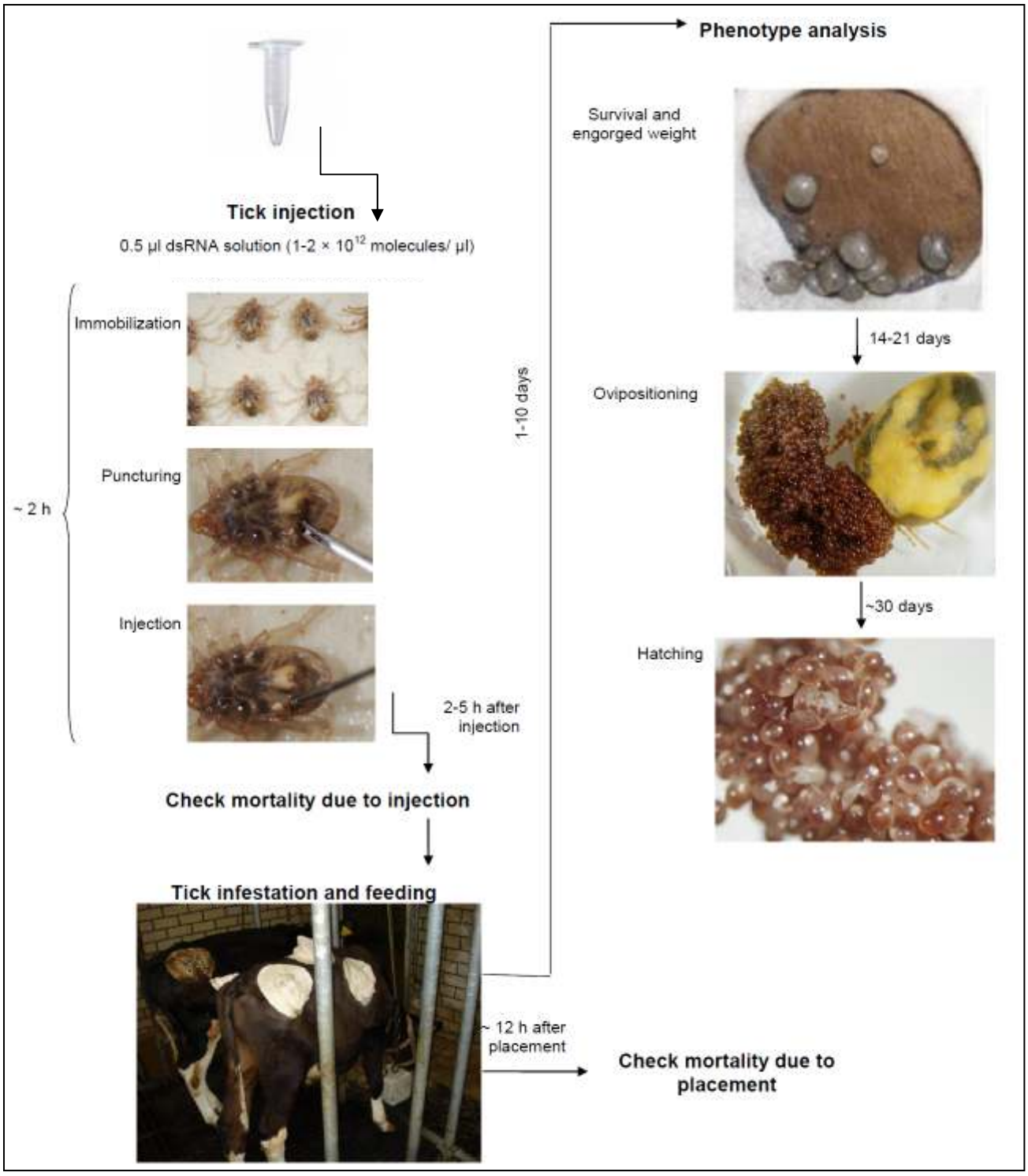

Figure 3.6. Typical *in vivo* RNAi experiment. Freshly molted *R. microplus* female ticks were mounted, with their ventral sides facing up, on double –sided tape. The exoskeleton surface was punctured with a 27G needle and the dsRNA or injection buffer (TE) control was injected into the lower quadrant of the ventral side with a 33G needle. After placement on the host and feeding, ticks survival, engorgement weight, oviposition and hatching rate were evaluated and compared between dsRNA-injected and control ticks.

Following injection the ticks were allowed to recover in an incubator at 27ºC with 95% relative humidity for 2-5 hours. Before infestation on cattle, each group's mortality due to injection was determined. Twelve hours after placement all dead ticks were removed from the patches and this mortality was recorded as mortality due to placement.

3.2.5 Phenotype analysis

The 5 groups of injected ticks (together with non-injected males) were placed on calf #03903, in 5 separate patches, respectively (Figure 3.6.). Ticks were monitored daily. Female ticks that detached upon repletion were collected, weighed and placed in 2 ml Eppendorf tubes with pierced lids at 27ºC with 95% relative humidity for oviposition. Several parameters including: egg weight; percentage female weight converted into egg weight; and the hatching rate were monitored and compared between dsRNA-injected and control ticks. The significance of the data of the weights of ticks after feeding and oviposited egg masses was calculated using Microsoft Excel and consisted of an unpaired *t*-test with unequal variances. Results were compared between the dsRNA and mock-injected ticks and P values of 0.05 or less were considered statistically significant.

3.2.6 Transovarial gene silencing

During transovarial gene silencing studies, groups of 5 engorged *R. microplus* female ticks were injected with 5 µl of each transcript's dsRNA alone or dsRNA solutions in the same combinations as stipulated above (Table 3.2.) or injection buffer alone (1-2 × 10^{12} molecules/ µl). Engorged female ticks were injected in their right spiracle plate (Figure 3.7), using the same methods as described above, within 6 h after dropping off the host. The injected females were stored individually in a 2 ml Eppendorf tubes (with pierced lids) at 27ºC with 95% relative humidity. Egg weight, percentage female weight converted into egg weight, and the hatching rate, were monitored and compared between dsRNA-injected and control ticks. Identical statistical analysis as described in section 3.2.5 was followed.

3.2.7 Gene silencing confirmation procedures

In order to assess gene silencing, RNA isolation followed by cDNA synthesis and semi-quantitative real-time PCR were performed. These three steps are discussed in the following sections.

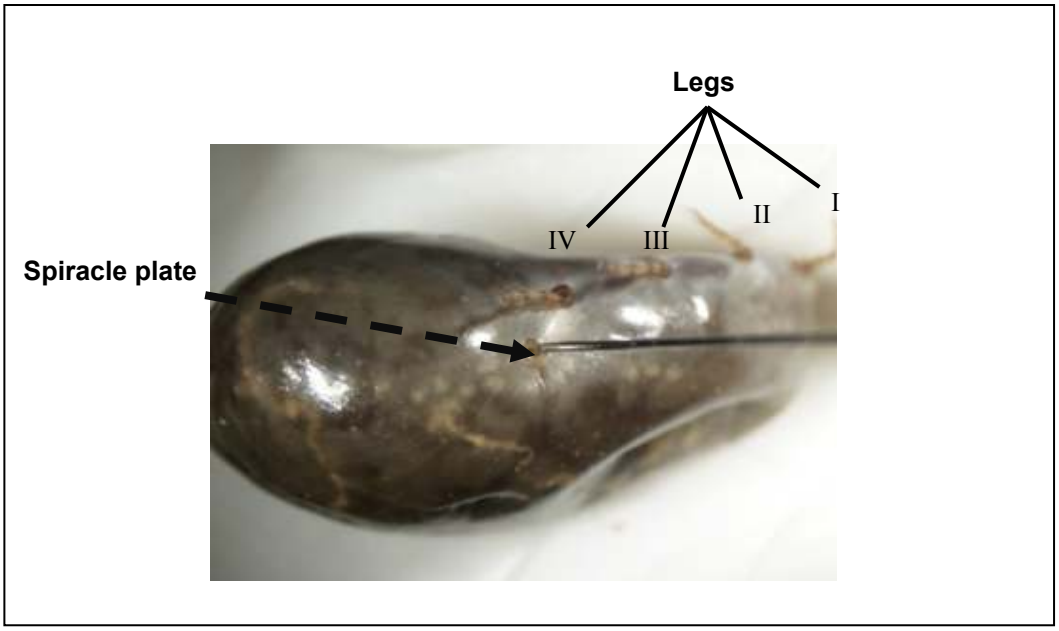

Figure 3.7. Injection for transovarial RNAi. A fully engorged tick was mounted in a piece of clay. The dsRNA of a target, or the injection buffer (TE) alone, were injected through the spiracle plate, using a 33G needle.

3.2.7.1 Tick dissection and RNA isolation

The viscera from partially fed females was dissected and used for subsequent organ collection and total RNA isolation. For each dsRNA injected or mock-injected group, 3 groups of 5 female ticks were dissected (biological repeats). During dissection each tick was submerged in a drop of autoclaved ice-cold phosphate buffer saline (PBS; pH 7.4) on a microscope slide. A cross-section between leg pair 2 and 3 was made, using a sterile scalpel blade. The tissues (specifically the salivary glands, midgut and ovaries) were collected from the ventral part of the body using watchmaker forceps, under a stereomicroscope. For each group of 5 ticks: the midguts were pooled into 1 ml TRIzol (Invitrogen™ life technologies, USA), while the salivary glands and ovaries were pooled

separately into 0.5 ml TRIzol. Dissected tissues, suspended in TRIzol, were kept on ice at all times. All samples were homogenised using physical passage through a large needle (20G) coupled to a sterile syringe, followed by passage through a smaller needle (27G).

Total RNA was isolated from the collected tissues utilising manufacturer's guidelines for TRIzol (Invitrogen™ life technologies, USA and (Chomczynski, 1993)). In brief, homogenised samples were centrifuged for 10 minutes at 12 000 x g (4ºC), the supernatants transferred to new 2 ml microcentrifuge tubes and the pellets (containing extracellular membranes, polysaccharides and high molecular weight DNA) discarded. Each sample was incubated for 5 minutes at room temperature. Chloroform was then added to each sample (100 µl per 500 µl TRIzol) and the samples were vigorously vortexed for 30 seconds and incubated at room temperature for 3 minutes. The reaction mixtures were centrifuged for 15 minutes at 12 000 x g (4ºC) to allow complete phase separation and the aqueous phases were transferred to new 2 ml microcentrifuge tubes containing 500 µl isopropanol. In order for thorough RNA precipitation to take place, the samples were gently mixed and incubated at room temperature for 15 minutes. Following centrifugation at 12 000 x g for 30 minutes (4ºC), the supernatants were removed and the pellets washed with 250 µl of 75% ethanol. Pellets were collected by centrifugation at 12 000 x g for 10 minutes (4ºC). The supernatant was removed and the RNA pellets were air-dried. Finally, the pellets were resuspended in RNase free water (DEPC-treated). RNA concentration as well as the $A_{260/280}$ and $A_{260/230}$ ratios was measured with the NanoDrop (Thermo-scientific, USA). All samples were stored at -80ºC.

3.2.7.2 DNase treatment of RNA and cDNA synthesis

For DNase treatment 3-20 µg RNA was combined with 10 µl DNase buffer (Fermentas GmbH, Germany), 1 µl per 1 µg of total RNA DNase (Fermentas GmbH, Germany), 1 µl RNase inhibitor (Fermentas GmbH, Germany) and RNase free water to a final volume of 100 µl. Each reaction mixture was incubated at 37ºC for 30 minutes. DNase activity was terminated with the addition of 10 µl of 10 mM EDTA and an incubation step of 10 minutes at 65ºC. To collect undigested RNA, 100 µl of phenol/chloroform/isoamyl alcohol was added to each DNase treated RNA sample, followed by vigorous vortexing

and brief centrifugation (at room temperature). The aqueous phase of each sample was transferred to a new microcentrifuge tube and one fifth of the volume sodium acetate (3 M, pH 5.2) was added and the reaction mixtures were mixed briefly. To allow precipitation, three volumes of ice-cold 100% ethanol were added; the reaction mixture was mixed and centrifuged for 30 minutes at 12 000 x g (4°C). The supernatants were discarded and precipitates were washed with 250 µl 70% ethanol and subsequent centrifugation. RNA pellets were air-dried and finally resuspended in RNase free water (DEPC-treated). The RNA concentration of each sample, as well as the $A_{260/280}$ and $A_{260/230}$ ratios was measured with the NanoDrop (Thermo-scientific, USA). All samples were stored at -80ºC.

cDNA was synthesised using the iScript™cDNA Synthesis Kit (Bio-Rad Laboratories, Inc., USA) exactly as described in section 2.2.5, using 500 ng RNA (of partially fed female: mid gut, salivary glands and ovaries). In order to analyse if all cDNA were intact, standard PCRs (as described in section 2.2.6) were conducted, using Elongation factor 1α gene specific primers.

3.2.7.3 Semi-quantitative real-time PCR

Semi-quantitative real-time PCR (semi-qPCR) was used to analyse the effect of RNAi on the transcription level of each target of interest. This fluorescence-based method allows for accurate relative quantification of starting amounts of nucleic acid targets (Heid *et al.*, 1996). Its combination of speed, sensitivity and specificity, together with its conceptual and practical simplicity, makes it the touchstone for nucleic acid quantification.

All detection methods used in real-time PCR involve fluorescence. Either sequence specific probes or non-specific labels are used as reporters. The reporter that was used in this study was the DNA interchelator, SYBR Green. This asymmetric cyanine (Figure 3.8.) has virtually no fluorescence when it is free in solution, but upon binding to dsDNA the dye emits a bright green fluorescence. In a PCR the fluorescence increases with the amount of dsDNA produced, therefore the dye provides a direct measurement of the double stranded product concentration during the course of the reaction. Advantages of

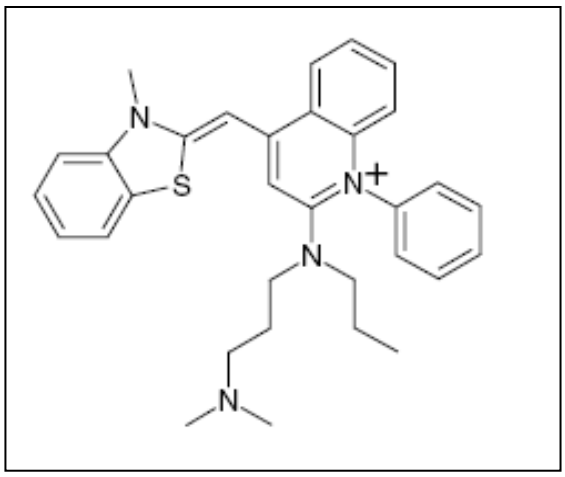

Figure 3.8. The asymmetric cyanine structure of SYBR Green.

SYBR Green include that it is inexpensive, easy to use and sensitive. However, it has the disadvantage of binding to any dsDNA including primer-dimers and non-specific amplified products (Kubista *et al.*, 2006).

With real-time PCR the initial copy number of a gene can be determined within a specific sample, since the real-time analysis of amplification allows for monitoring of the entire response curve. Based on the amount of initial template, different samples' response curves separate in the exponential phase of the reaction (Figure 3.9). A fixed fluorescence threshold is set above the base line (background) and for each sample the

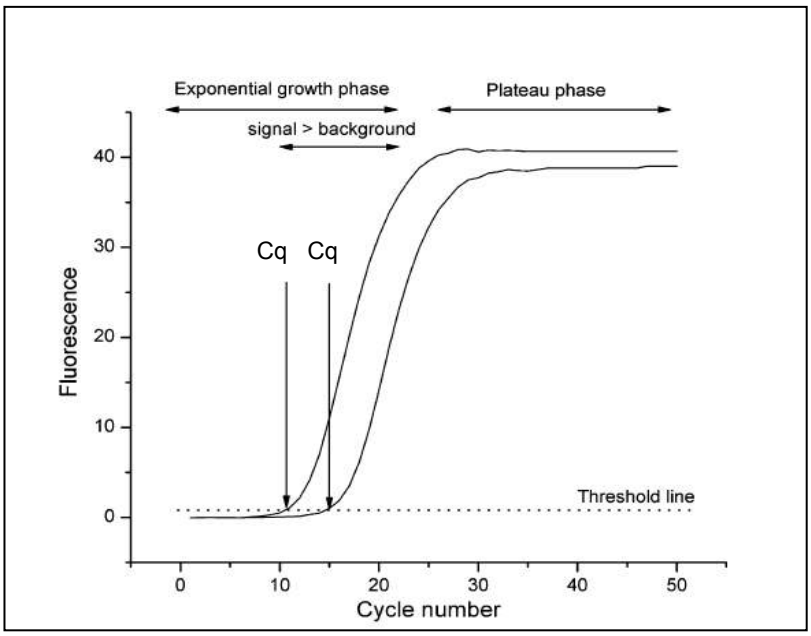

Figure 3.9. Real-time PCR response curves. A threshold base line is set sufficiently above background and the number of cycles to reach the threshold (Cq) are registered (adapted from Kubista *et al.*, 2006).

amount of cycles to reach the threshold (Cq) is determined (Heid *et al.*, 1996). This real-time measurement enables the calculation of amplification efficiencies for individual transcripts. With the use of appropriate software, it can be used to determine and compare the input copy number of any transcript within different samples (e.g. control samples vs. silenced samples).

Gene specific primer design for semi-qPCR:

The first step in order to conduct real-time PCR was to design a set of gene specific primers for each target of interest. The primers were selected to amplify a ±100 bp part within each gene, which is outside of the region that was targeted by the dsRNA, to circumvent the re-amplification of any unprocessed dsRNA. The annealing temperature of each primer was calculated (as described in section 2.2.3) and the primers were synthesised by Inqaba Biotec (Pretoria, RSA). The primers were dissolved in double distilled deionised water and stock solutions, with final concentrations 100 µM, were prepared. All primers were validated by means of a standard PCR and agarose gel electrophoresis.

Conduction of semi-qPCR:

Once gene specific primers were designed and synthesised and optimal conditions were established, semi-qPCR was conducted to determine the transcriptional level of each target in the control (TE) and its injected group (e.g. BmMP5 in group A).

All real-time PCRs were performed in 10 µl reaction volumes (in 384 well plates) using the KAPA™Syber® FAST qPCR kit (KAPABiosystems, USA) in accordance with the manufacturer's protocol, on the LightCycler® 480 (Roche Diagnostics, Germany). The cDNA, synthesised from the dissected tissues, was diluted 1/10 times with water and served as template. One microlitre template was added to 15 pmol of each primer, 1 µl water and 5 µl KAPA™Syber® FAST qPCR 2x Master Mix, which contains a novel DNA polymerase specifically evolved for qPCR using Syber® Green I dye chemistry. For each transcript a standard curve was constructed, using a dilution series (1/5; 1/10; 1/20; 1/50 and 1/100 dilutions) of *R. microplus* mixed lifestages as template. For each primer set a negative control was included: a reaction which lacked template. In order to

correct for run-to-run differences an intra-run calibrator was incorporated in each run. For each reaction, three technical repeats were performed. Once set up a plate was subjected to the LightCycler® 480, in which the enzyme was activated and the cDNA was denatured at 95ºC for 5 minutes. Amplification consisted of 45 cycles of cDNA denaturing (95ºC, 3 s), annealing (55ºC, 7 s) and extension (72ºC, 4 s). To discriminate between primer-dimers and specific product all amplified products were directly subjected to melting curve analysis, which included a denaturing (95ºC, 30 s) and annealing (55ºC, 30 s) step followed by a continuous acquisition mode (from 95ºC, with a ramp rate of 0.1ºC/s). Finally, all plates were cooled at 40ºC for 30 s. Each reaction's Cq value and melting temperature (Tm, of the product) were calculated with the LightCycler® 480 software.

Reference gene validation:

To assure that all real-time PCR data was reliable, an essential normalisation step was conducted, using internal reference genes (see following section). However, before any gene was used for normalisation, it was first experimentally validated for the particular tissues and experimental design used during this study (Guénin *et al.*, 2009). Nine candidate reference genes (Table 3.3.), which have been identified by Dr. A. Nijhof

Table 3.3. Characteristics of the nine candidate reference genes

Symbol	Gene name	Function	*R. microplus* GenBank accession number
ACTB	Beta actin	Cytoskeletal structural protein	AY255624
BTUB	Beta tubulin	Component of microtubules	CK179480
ELF1A	Elongation factor 1-alpha	Component of the eukaryotic translational apparatus	EW679365
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	Oxireductase in glycolysis and gluconeogenesis	CK180824
GST	Gluthathione S-transferase	Detoxification of endogenous components	CV456312
H3F3A	H3 Histone family 3A	Involved in structure of chromatin	CV442167
PPIA	Cyclophilin	Facilitate protein folding	CV445080
RPL4	Ribosomal protein L4	Structural component of the large 60S ribosomal unit	CV447629
TBP	TATA box binding protein	Transcription factor	CV453818

(personal communication) were validated, using real-time PCR and the software program qBasePlus (version 1.2) (Hellemans *et al.*, 2007).

Briefly, for each candidate reference gene a set of gene-specific primers, which amplifies a product of ±100 bp, was designed and synthesised by Inqaba Biotec (Pretoria, RSA) (Table 3.4.). Real-time PCRs were conducted exactly as described above, using 1/5 times dilution of cDNA from the TE group as template. Once the reactions were amplified and the LightCycler® 480 software has calculated the Cq values for each gene in each tissue of interest, the crude (non-normalised) data was exported to qBasePlus.

qBase software includes several algorithms which calculate for each gene the pair wise variation with all other genes, in terms of the standard deviation of their logarithmically transformed expression ratios. In addition, the program also calculates a coefficient of variance (CV) of the normalised reference gene quantities, which gives a direct indication of how stable a gene is expressed in a particular sample range. With the use of these functions the stability of each candidate reference gene was calculated. Genes with CV values lower than 25% were selected and used for normalisation.

Table 3.4. Characteristics of the gene specific primers, of the nine candidate reference genes, used in semi-quantitative real-time PCR.

Primer Name	Primer sequence (5' → 3')	Tm (°C)
ACTBfw	CCC ATC TAC GAA GGT TAC GCC	61.78
ACTBrv	CGC ACG ATT TCA CGC TCA G	58.83
B86/Ra86fw	CGT CCC GAC TTG ACC TGC	60.52
Bm86/Ra86rv	AGG AGC GGC TGA ACA GTT TG	59.35
BTUBfw	AAC ATG GTG CCC TTC CCA CG	61.40
BTUBrv	GCA GCC ATC ATG TTC TTT GC	57.30
ELF1Afw	CGT CTA CAA GAT TGG TGG CAT T	58.39
ELF1Arv	CTC AGT GGT CAG GTT GGC AG	61.40
GAPDHfw	AGT CCA CCG GCG TCT TCC TCA	63.73
GAPDHrv	GTG TGG TTC ACA CCC ATC ACA A	60.25
H3F3Afw	AAG CAG ACC GCC CGT AAG T	58.83
H3F3Arv	GTA ACG ACG GAT CTC CCT GAG	61.78
PPIAfw	CTG GGA CGG ATA GTA ATT GAG C	60.25
PPIArv	ATG AAG TTG GGG ATG ACG C	56.67
RPL4fw	AGG TTC CCC TGG TGG TGA G	60.98
RPL4rv	GTT CCT CAT CTT TCC CTT GCC	59.82
TBPfw	CTT GTC CTC ACA CAC AGC CAG TT	62.43
TBPrv	GTG AGC ACG ACT TTT CCA GAT AC	60.65

Real-time PCR data and statistical analysis:

For real-time PCR data analysis all Cq-values (calculated by the LightCycler®480 software) were exported to qBasePlus. This flexible and open source program allowed for the processing of the raw data (of the multiple runs) into normalised and calibrated relative quantities.

qBasePlus is composed of two models: the 'q-Base Browser' which was used for managing and archiving the data and the 'qBase Analyzer' which was used for the processing of the raw data. The latter process involved several consecutive steps: initialisation; reviewing of sample and gene annotation; reference gene selection (as described above); raw data control; sample organisation and selection; calculation of amplification efficiencies; calculation of relative quantities and evaluation of the results. The conversion of the quantification cycle values (Cq) into normalised relative quantities (NRQs) was done with the use of a modified delta-delta Cq formula:

$$NRQ = \frac{E_{goi}^{\Delta Ct,goi}}{\sqrt[f]{\prod_{o}^{f} E_{ref_o}^{\Delta Ct,ref_o}}}$$

This model constitutes an improvement over the classic delta-delta Cq (ΔΔCq) method, since it accounts for differences in PCR efficiencies (E) between the gene of interest (goi) and the reference genes (ref), as well as for multiple (f) reference genes.

Finally the processed data (with standard error values) were exported to Microsoft Excel. The significance of each transcript's percentage silencing, within each separate biological repeat was determined using a two-tailed *t*-distribution with unequal variances. To test the significance of each transcript's percentage silencing overall (thus all three biological replicates merged), an unpaired *t*-test with unequal variances was performed using Microsoft Excel. The results were compared between the dsRNA and mock-injected tissues and P values of 0.05 or less were considered statistically significant.

Note that the normalised transcriptional level of each transcript was only determined within the tissue type it is known to be highly expressed in. No subsequent real-time analysis was performed for BmMP3 and As70, since both transcripts had very low expression levels. No real-time analysis was performed on eggs obtained from the transovarial studies, since no significant phenotype was obtained for the transcripts which are expressed in the eggs.

3.2.8 Differential transcriptional response investigation

Both to assess if any non-specific silencing occurred and to investigate the differential transcriptional responses between these metzincins, the expression level of each transcript within all the different dsRNA-injected groups (in all the relevant tissue types) was determined. This was done by means of integrated semi-quantitative real-time PCRs. All reactions were conducted exactly as described above, in section 3.2.7.3, utilising the same sets of gene-specific primers (Table 3.5 and 3.6) used during the silencing confirmation experiments. All the crude (non-normalised) Cq-data was exported from the LightCycler® 480 software to qBasePlus, where it was normalised to the selected reference genes. The processed data (with standard error values) was finally exported to Microsoft Excel, where the normalised levels of the transcripts within all the different dsRNA-injected groups were compared with their normalised levels in the mock-injected (control) group. The significance of any non-specific silencing and of any transcriptional up regulation was determined by unpaired *t*-tests, with unequal variances using, Microsoft Excel. P values of 0.05 or less were considered statistically significant.

3.2.9 Metalloprotease assay

In order to determine the subsequent effect of RNAi on the protein level, the overall metalloprotease activity was determined for the investigated tissues. This was done with the use of the Protease Fluorescent Detection Kit (Sigma-Aldrich, USA), 1,10-Phenanthroline (Sigma-Aldrich, USA) and the Fluoroskan Ascent FL (ThermoLabsystems, USA) fluorescent meter. The fluorescence-based assay could only provide the total protease activity of a sample. Therefore, to single out

metalloprotease activity, the assay was also performed using samples treated with the metalloprotease inhibitor, 1,10-Phenanthroline. The metalloprotease activities were then indirectly calculated from the two assays' results combined.

The Protease Fluorescent Detection Kit uses a modified method of the Fluorescein Isothiocyanate (FITC)-labeled casein assay for proteolytic enzymes, first described by Twining (1984). In essence, to detect protease activity the kit makes use of casein labeled with FITC, as substrate. In the presence of proteases the FITC-labeled casein is cleaved into smaller fragments. These fragments do not precipitate under acidic conditions. Therefore, after incubation of a protease sample and the FITC-labeled casein substrate; an acidification step (with trichloroacetic acid (TCA)); and a centrifugation step the fluorescence of the FITC-labeled fragments can be determined, giving a direct measure of protease activity.

For the assay 5 partially fed female *R. microplus* ticks, from each dsRNA-injected group and the mock-injected group, were dissected (as described in section 3.2.7.1). For each group of 5 ticks: the salivary glands, midguts and ovaries were pooled into 50 µl ice-cold PBS (pH 7.4). To obtain the total protein from each group's dissected tissues the samples were sonified, with a Branson Model B-30 sonifier (Branson Sonic Power Co.), set at 10 pulses at 20% duty cycles at an output control of 2. The samples were centrifuged (at 16 000 x g, 4ºC, 20 minutes) and the supernatant - containing the total protein – were aliquoted into new microcentrifuge tubes.

The protein concentration of each sample was determined, using the NanoDrop ND-1000. From each sample two 10 µl aliquots were transferred to separate microcentrifuge tubes. Ten millimolar 1,10-Phenanthroline was added to one of the tubes and the reaction mixture was mixed well by pipetting. Thereafter, 20 µl Incubation Buffer and 20 µl FITC-casein substrate were added to both the tube treated with 1,10-Phenanthroline and untreated tube. Each tube was mixed gently and was incubated at 37ºC in the dark for 60 minutes. After incubation 150 µl 0.6 N TCA solution was added to each tube. The reaction mixtures were mixed gently and incubated at 37ºC in the dark for 30 minutes. The samples were centrifuged at 10 000 x g at room temperature for 10 minutes. Two microlitre of each sample's supernatant (containing the acid soluble FITC labeled fragments), together with 200 µl Assay buffer, were transferred to a well of a 96 well plate. The plate was shaken well and the fluorescence intensity of each

sample was determined at a 485 nm excitation and 535 nm emission wavelength. A blank sample, which was prepared from 10 µl water instead of 10 µl protein sample, was included. A standard curve was constructed using a dilution series (1/5; 1/10; 1/20; and 1/40) of 20 µg/ml Trypsin. Due to limited quantities of extracted total protein, all reactions were done only in duplicate.

The data calculated by Fluoroskan Ascent FL software were exported to Microsoft Excel. The total protease activity was determined in terms of concentration, in both the 1,10-Phenanthroline treated and untreated samples, using the Trypsin standard curve. The concentration of only the metalloproteases present in each sample was then calculated indirectly, by subtracting the concentration of the treated sample (all proteases except metalloproteases) from the untreated (total proteases). The standard deviations of the mean, for each two technical repeats were determined, using Microsoft Excel. Finally, the metalloprotease activities within the different tissues of each dsRNA-injected group were compared to the metalloprotease activities within the different tissues of the control group.

3.3 Results and Discussion

3.3.1 RNA interference

To assess the function and the significance of the roles of the 5 reprolysin and 3 astacin *R. microplus* metzincins in tick development, feeding and reproduction, the phenotypical effect upon silencing of these genes was studied. Gene silencing was performed by means of *in vivo* RNAi in both partially fed and fully engorged *R. microplus* females.

3.3.1.1 dsRNA gene specific primer design and T7-DNA template synthesis

A set of gene specific primers (forward and reverse), with incorporated T7 promoter sequences at the 5' ends, was designed for each of the 5 reprolysins and 3 astacins *R. microplus* transcripts. The characteristics of these primers are summarised in Table 3.5 and 3.6 below. The region of each transcript's nucleic acid sequence which was targeted for silencing, was randomly selected. The sizes of these regions, of the different transcripts, ranged between 120 bp and 635 bp.

Table 3.5. Characteristics of the T7 incorporated gene specific primers used in dsRNA synthesis of the 5 reprolysin-like transcripts.

Primer Name	Primer sequence (5' → 3')	Tm (°C)
iBmMP1fw	TAA TAC GAC TCA CTA TAG GGA GAT ACC AGA AGG CGA GAG GTC AAT G	73.9
iBmMP1rv	TAA TAC GAC TCA CTA TAG GGA GAA CAA GAA GCA GGG CAT CGT G	73.3
iBmMP2fw	TAA TAC GAC TCA CTA TAG GGA GAG CAC GCT TGT GGT TCA CGA A	73.3
iBmMP2rv	TAA TAC GAC TCA CTA TAG GGA GAC GCA TCA GGA TCG CCA TAT AGA T	73.0
iBmMP3fw	TAA TAC GAC TCA CTA TAG GGA GAT AAG TGA AGA CAT AAC GCT GAA C	71.2
iBmMP3rv	TAA TAC GAC TCA CTA TAG GGA GAG CTC GTG GTC GTC GTA AA	72.4
iBmMP4fw	TAA TAC GAC TCA CTA TAG GGA GAC CAC GCC ACA GGA TTG AAC	73.3
iBmMP4rv	TAA TAC GAC TCA CTA TAG GGA GAT ATC GTT GCC ATT TCT GTA AGG G	72.1
iBmMP5fw	TAA TAC GAC TCA CTA TAG GGA GAA TTG CGG AAC ATC ATT ACC ATA A	70.3
iBmMP5rv	TAA TAC GAC TCA CTA TAG GGA GAG GTT TCT AAG GGC ACT GTA TCG T	73.0

Table 3.6. Characteristics of the T7 incorporated gene specific primers used in dsRNA synthesis of the 3 astacin-like transcripts.

Primer Name	Primer sequence (5' → 3')	Tm (°C)
iAs5051fw	TAA TAC GAC TCA CTA TAG GCG TGA CGG AGC CGA ATA GAC TGT	73.3
iAs5051rv	TAA TAC GAC TCA CTA TAG GAA TGG CGT CAG CAG TCG GTT C	72.5
iAs8770fw	TAA TAC GAC TCA CTA TAG GAC CTG CCT GCG ATT TGT TGA	70.5
iAs8770rv	TAA TAC GAC TCA CTA TAG GCG CAT CCG CAG AAT ACA GCA T	71.5
iAsContigfw	TAA TAC GAC TCA CTA TAG GGA GAG TGT TTT CTG GCT CCA CGT T	72.3
iAsContigrv	TAA TAC GAC TCA CTA TAG GGA GAA TGC GGT TAC TAA GAA GGA CGA G	73.0

For the production of dsRNAs, double stranded DNA templates with flanking T7 regions needed to be synthesised. This was done by means of PCR amplification, utilising the latter primers and single stranded cDNA (synthesised from *R. microplus* mixed lifestages total RNA) as initial template. Upon analysis of the PCR products on 2% Agarose/TAE/EtBr gels (Figure 3.10) it was clear that for each transcript a single band of correct size was obtained, indicating that all primers annealed specifically and correctly to their coding sequences. Products slightly larger than the expected sizes were obtained; however, this can be explained by the incorporation of the two T7 sites (± 50 bp in total) flanking the sequences.

Three identical PCR reactions of each transcript were pooled together. The PCR-product was purified and the concentration determined. For each of the 5 reprolysin and 3 astacin metzincins sufficient amounts of T7 incorporated DNA was obtained, to serve as template for subsequent dsRNA synthesis.

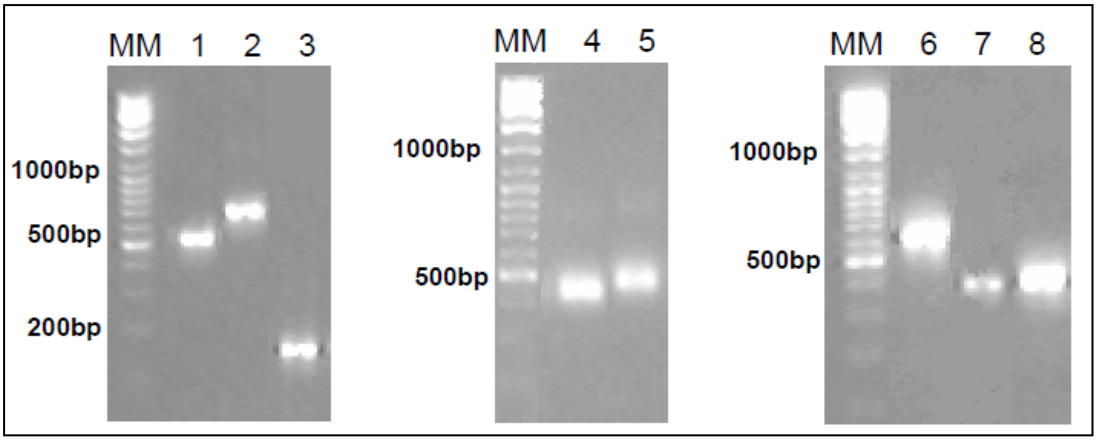

Figure 3.10. Agarose gel electrophoresis of the T7 incorporated DNA templates of the 5 reprolysin and 3 astacin metzincins. The first lanes correspond to molecular mass markers (MM). The numbered lanes correspond to: (1) BmMP1, (2) BmMP2, (3) BmMP3, (4) BmMP4, (5) BmMP5, (6) As51, (7) As71 and (8) AsC.

3.3.1.2 dsRNA synthesis

For each of the 8 *R. microplus* metzincins, complementary RNA strands were synthesised from their respective T7 incorporated DNA template. After the transcription reactions, the resulting paired RNA strands were annealed to form dsRNA. For accurate RNA concentration determination all remaining DNA template was removed by DNase treatment and purification. Furthermore, to assure that all remaining ssRNA were removed and that dsRNA was in fact synthesised, the samples were subjected to RNase treatment. Upon analysis of the products on 2% Agarose/TAE/EtBr gels it was clear that for all transcripts a single distinct dsRNA band correlating to the expected size was obtained (Figure 3.11.), confirming that intact dsRNA was successfully synthesised for each transcript.

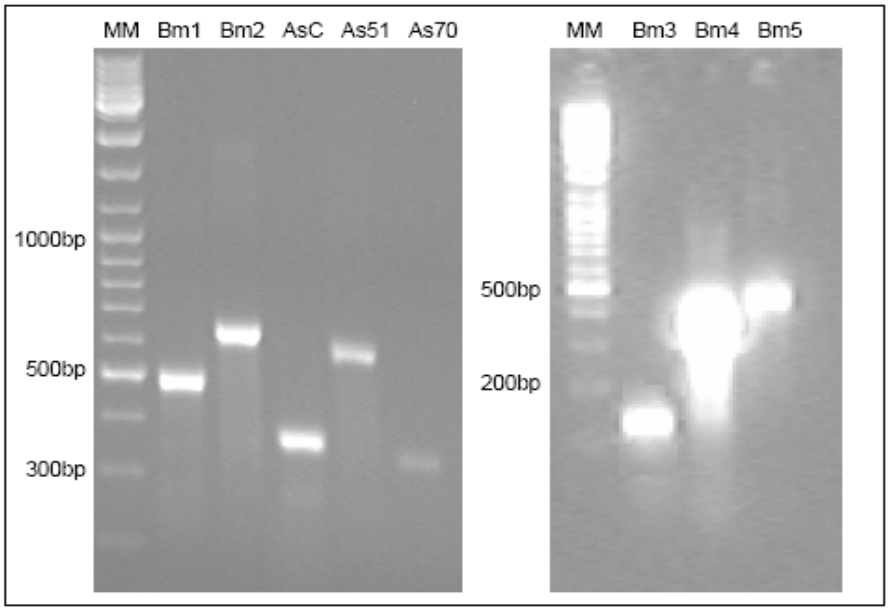

Figure 3.11. The synthesised dsRNA of the 5 reprolysin and 3 astacin metzincin resolved on 2% agarose gels. The first lanes correspond to molecular mass markers. For Bm1, Bm2, AsC, As51 and As70 a 1/50 dilution of each sample was visualised, for Bm3 and Bm5 a 1/10 dilution of each sample was visualised and for Bm4 undiluted sample was visualised.

From the concentration data the number of molecules of dsRNA, present in each purified sample, was quantified (Table 3.7). Since a higher number of molecules will in effect yield a greater silencing effect; the dosage of dsRNA was based on the number of molecules administered rather than on concentration, to assure the variation between dsRNAs of different sizes (therefore different molecular weights) is minimal.

Table 3.7. The concentration (ng/µl) and the number of molecules (per µl) of synthesised dsRNA of each metzincin transcript.

	Concentration dsRNA (ng/ul)	molecules / ul
BmMP1	2370	4.58E+12
BmMP2	3546	5.25E+12
BmMP3	2724.2	1.78E+13
BmMP4	1352.6	3.36E+12
BmMP5	1767.3	3.69E+12
As51	4253.6	6.56E+12
As70	4540.2	1.36E+13
AsC	4086.6	1.05E+13

3.3.1.3 Tick injection and phenotype analysis

For this study *in vivo* gene silencing was performed, by injecting gene-specific double-stranded RNA (dsRNA) into freshly molted unfed adult females (A) and engorged females for transovarial silencing (B) to assess a possible phenotype of these metalloproteases in feeding and oviposited eggs of female *R. microplus*, respectively.

A.) Silencing in freshly molted adult females (silencing prior to feeding)

Five groups consisting of 80 freshly molted *R. microplus* females were injected with 0.5 µl injection buffer (containing 1-2 × 10^{12} molecules/ µl of each transcript) as follows: BmMP4 and BmMP5 (Group A); BmMP1, BmMP2 and BmMP3 (Group B); As51 and As70 (Group C); AsContig (Group D) and injection buffer alone (TE/ Control Group). The method of injection was successful since an average of 72.8 (9% overall mortality) females were alive in each group 2-4 hours following injection (Table 3.8.). These ticks

Table 3.8. Tick number parameters monitored throughout the RNAi study. (ffF: fully fed females)

Group	TE	A	B	C	D
Ticks injected	80	80	80	80	80
Mortality after injection	12	12	6	1	5
Mortality after injection (%)	15.00%	15.00%	7.50%	1.25%	6.25%
Ticks placed on cattle	68	68	74	79	75
Ticks removed for dissection	20	20	20	20	20
ffF collected for ovipositioning	21	21	34	44	25
Total removed	41	41	54	64	45
Mortality due to placement	27	27	20	15	30
Number of ffF that laid no eggs	1	2	5	8	0

were subsequently fed together with an excess of *R. microplus* males until the females became replete or for a maximum of 14 days. Several factors, not pertaining to the gene silencing, were checked to ensure that a phenotype obtained would reflect a true positive.

Tick weight after engorgement (or manual removal), mortality rate after placement, egg mass and oviposition efficiency (the percentage of fully fed female (ffF) weight converted into egg mass weight) were monitored and are presented in Table 3.9.

Table 3.9. Tick engorgement weight, egg mass weight and oviposition efficiency of double-stranded RNA (dsRNA)-injected *R. microplus* ticks, injected as freshly molted females. (ffF: fully fed females)

Group	Number of ticks	Average ffF weight[a] (mg)	Average egg weight[b] (mg)	Average ffF to egg[c] (Oviposition efficiency)
TE	68	328±49	152±63	46%
A	68	364±71	129±54	37%
B	74	352±58	93±56**	27%**
C	79	341±73	89±72**	26%**
D	75	313±63	111±38*	37%

[a] Ticks which completed feeding and those removed 14 days after mock- or dsRNA-injection were weighed, the average ± SD weight was calculated and compared between dsRNA and mock-injected control ticks, using the Student's *t*-test with unequal variance.

[b] The egg mass oviposited by each tick was weighed individually and compared between dsRNA- and mock-injected control ticks, using the Student's *t*-test with unequal variance (*P<0.05, **P<0.01).

[c] The percentage fully fed female weight converted into egg mass was calculated for each tick individually and compared between dsRNA- and mock-injected control ticks, using the Student's *t*-test with unequal variance (*P<0.05, **P<0.01).

As shown in Table 3.9, simultaneous injection of BmMP1, BmMP2 and BmMP3 (Group B) and simultaneous injection of As51 and As70 (Group C), respectively, resulted in significant decreased oviposited egg masses and oviposition efficiencies compared with the control group (P<0.01). Although slightly less significant (P<0.05), injection of AsC (Group D) also resulted in a decrease of egg mass oviposited, no phenotype in regard with the oviposition efficiency was observed. The hatching rates of all four dsRNA-injected groups were monitored and compared to the mock-injected group but no significant change was detected.

B.) Transovarial silencing (silencing in fully engorged females)

For the transovarial gene silencing studies, engorged *R. microplus* females were injected with 5 µl of BmMP1, 2, 3, 4, 5 or As51, 70, Contig dsRNA individually, or dsRNA solutions in the same combinations as mentioned above (Group TE, A, B, C and D), or injection buffer alone in the right spiracular plate, within 6 h after dropping off the host. The method of injection was successful since no ticks died due to injection. The injected female ticks were stored individually and several parameters including: egg weight; oviposition efficiency (percentage female weight converted into egg weight); and the hatching rate were monitored and compared between dsRNA-injected and control ticks (Table 3.10).

Table 3.10. Transovarial silencing results. Egg mass weight, oviposition efficiency and hatching rate of double-stranded RNA (dsRNA)-injected *R. microplus* ticks, injected as fully engorged females are indicated. All transcripts were injected individually, but only results deviating from TE is given. Data representative only of ticks that did in fact lay eggs. (ffF: fully fed females)

Group	Average egg weight[a] (mg)	Average ffF to egg[b] (Oviposition efficiency)	Avg ffF to egg as % of TE	Hatching rate[c]
TE	153.10 ± 44.45	44.82%	100%	>90%
A	106.50 ± 50.34	21.88%*	48.8%*	>90%
B	136.00 ± 2.83	34.51%	76.9%	>90%
C	89.75 ± 47.72	26.10%	58.2%	>90%
D	110.40 ± 43.37	29.02%	64.7%	>90%
Individual Transcripts				
BmMP4	146.60 ± 45.62	38.26%	85.4%	>90%
BmMP5	166.40 ± 12.30	38.12%	85.1%	>90%

[a] The egg mass oviposited by each tick was weighed individually and compared between dsRNA- and mock-injected control ticks, using the Student's *t*-test with unequal variance.

[b] The percentage fully fed female weight converted into egg mass was calculated for each tick individually and compared between dsRNA- and mock-injected control ticks, using the Student's *t*-test with unequal variance (*P<0.01).

[c] The hatching rate was determined 6 weeks post oviposition and compared between control injected and dsRNA-injected ticks using the Student's *t*-test with unequal variance

From the results it is evident that simultaneous injection of BmMP4 and BmMP5 (Group A) yielded a significant phenotype: the oviposition rate was significantly reduced (P<0.01) when compared with the control group. What is noteworthy is the fact that the individual silencing of either BmMP4 or BmMP5 failed to produce a significant

phenotype. Therefore it can be postulated that a synergistic relationship between these two transcripts exist.

No significant phenotype regarding the hatching rate was observed, therefore it can be reasoned that the silencing of these metzincins had no effect on the processes of larvae hatching and development and extracellular coat degradation (hatching). However, due to the limited number of ticks used during this study, the gene silencing was not confirmed by subsequent real-time PCR. Thus it remains indefinite if silencing took place in the developing eggs. To obtain reliable data, this experiment needs to be repeated.

By integrating the results of both *in vivo* gene silencing studies (as well as the expression profile studies in Chapter 2), some postulations regarding the importance of the different metzicins throughout tick feeding and reproduction can be made.

The mechanism that impaired the ticks to successfully produce ample amounts of eggs is unknown. However, it is known that it is mandatory for adult female ixodid ticks to have successful blood meals to produce eggs (Sonenshine, 1991). With the completion of feeding and digestion a hormonal signal triggers vitellogenesis, a vital process in oogenesis which produce the yolk nutrient for the developing larvae (Sonenshine, 1991). Therefore, it can be speculated that silencing of the reprolysin and astacin metalloproteases (within the salivary glands and the midgut) could have indirectly impaired oogenesis by directly impairing blood feeding and digestion, since these metalloproteases are hypothesised to be involved in digestion of the extracellular matrix (ECM) and other aniti-haemostatic components. Since the astacin transcripts are also present in the ovaries and are known to be involved in vital functions such as growth factor activation (Stöcker *et al.*, 1993) and embryo development (dorso-ventral patterning) (Shimell *et al.*, 1991), it can be speculated that the silencing of these enzymes could have directly affected oogenesis.

Due to limited amounts of sequence data available, off-target gene silencing cannot be excluded. To our knowledge the sequences of the dsRNAs used in this study do not contain any significant overlap with any other known *R. microplus* sequences. More profound knowledge of ixodid tick sequences, such as tick specific genome projects and

functional studies, will greatly improve and create efficient resources for the identification and characterisation of more effective tick protective antigens.

3.3.2 Semi-quantitative real-time PCR studies

3.3.2.1 Reference gene validation

Normalisation of relative quantities with reference genes relies on the assumption that the reference genes are stably expressed in the particular test sample range. Therefore, the first step, before any semi-quantitative real-time PCRs could be performed, was to validate and select suitable reference genes in *R. microplus*. Nine candidate genes (Table 3.3) were validated. The expression level of each gene was determined in the salivary glands, ovaries and midgut of the control (mock-injected) group by means of real-time PCR. The data was exported to qBasePlus, where the stability of each gene was calculated and given in terms of a coefficient of variation (CV). Ideally, a reference gene should display the same expression level across all samples after normalisation (thus CV= 0). Only two out of the nine genes, elongation factor 1 alpha (ELF1α) and cyclophilin (PPIA), displayed CV values smaller than 25% (CV< 0.2) which falls into the acceptable norm (Hellemans *et al.*, 2007), and were therefore selected and used as reference genes for normalisation in subsequent analysis.

Using multiple stably expressed reference genes is currently considered to be the gold standard for normalisation (Bustin *et al.*, 2009; Guénin *et al.*, 2009). Not only do multiple reference genes produce more reliable data, but it also permits an evaluation of the stability of these genes (Hellemans *et al.*, 2007).

3.3.2.2 Percentage silencing confirmation

To confirm that the observed phenotypes were in fact due to gene-specific silencing, the percentage silencing of each transcript within its respective dsRNA-injected group was determined. This was done by means of semi-quantitative real-time PCRs, using gene

specific primers designed to amplify regions outside of the regions used for dsRNA synthesis (Table 3.11.).

Expression levels of each transcript were only determined in tissues previously identified to express the transcript in question. With the use of qBasePlus all raw (unprocessed) Cq-values were normalised against both ELF1α and PPIA and the normalised expression level of each transcript within its silenced group was compared to its normalised expression level in the control group (Figure 3.12. and 3.13.).

The normalised transcript levels of the reprolysin-like transcripts (BmMP1, BmMP2, BmMp4 and BmMp5) (Figure 3.12.) were significantly reduced in the salivary glands of the respective groups ($P<0.01$). For BmMP4 and BmMP5 the normalised transcript levels were reduced with 77.15% and 97.42%, respectively in the combined BmMP4/BmMP5/-dsRNA-injected group (Group A). The normalised levels of BmMP1 and BmMP2 were reduced with 64.54% and 97.36%, respectively, in the combined BmMP1/BmMP2/BmMP3-dsRNA-injected group (Group B). The normalised transcript level of As51 was significantly reduced, in both the midgut with 55.82% ($P<0.01$), the ovaries with 80.42% ($P<0.01$) and the salivary glands with 83.75% ($P<0.01$) in the combined As51/As70-dsRNA-injected group (Group C), whereas AsC was only

Table 3.11. Characteristics of the gene specific primers used in semi-quantitative real-time PCR.

Primer Name	Primer Seq 5'-3'	Tm
qBmMP1fw	GCA AGA GAG ATC AAC CGA AAG	57.87
qBmMP1rv	AAA GAG AAG TTT GTC CGC AAG TA	57.08
qBmMP2fw	AAT CCG TGC GCT ATT GTT GCT AC	60.65
qBmMP2rv	GCC AAT CCA TCG TGA ATG CTA AC	60.65
qBmMP3fw	TTA TGA CAG GAT CGG CTA ATC TA	57.08
qBmMP3rv	TTA TAT GAA CCA TGT ACG GCT CT	57.08
qBmMP4fw	AAC TGA CGC TGA ACC TCA GAA AG	60.65
qBmMP4rv	CCG TTG AAG AAA TGT GTC ACT TC	58.87
qBmMP5fw	ACG GAA CGA AAT GAC ACA TAT CA	57.08
qBmMP5rv	GCA GAT CCA ACA AAG GCA TAA C	58.39
qAs5051fw	AAC CGA CTG CTG ACG CCA TT	59.35
qAs5051rv	CTC TGT CAA GTG ACT GCC GTC CT	64.21
qAsContigfw	TTG AGG TGA CTG CCA TTC TTC GC	62.43
qAsContigrv	AGA ACA GGT TGC TGA CGC CCT TC	64.21
qAs70Nqfw	TTG GTG TGG TGT AAG TTG ACC CT	60.65
qAs70Nqrv	GAA GTG ATT TGC CTG CCG TTT A	58.39

significantly silenced within the ovaries with 62.55% (P<0.01) of the AsC-dsRNA-injected group (Group D), compared to the normalised levels in the mock-injected group (Figure 3.13.). These results are comparable with previous semi-quantitative measurements of the gene silencing effect by dsRNA-injection in *R. microplus* (Nijhof *et al.*, 2007).

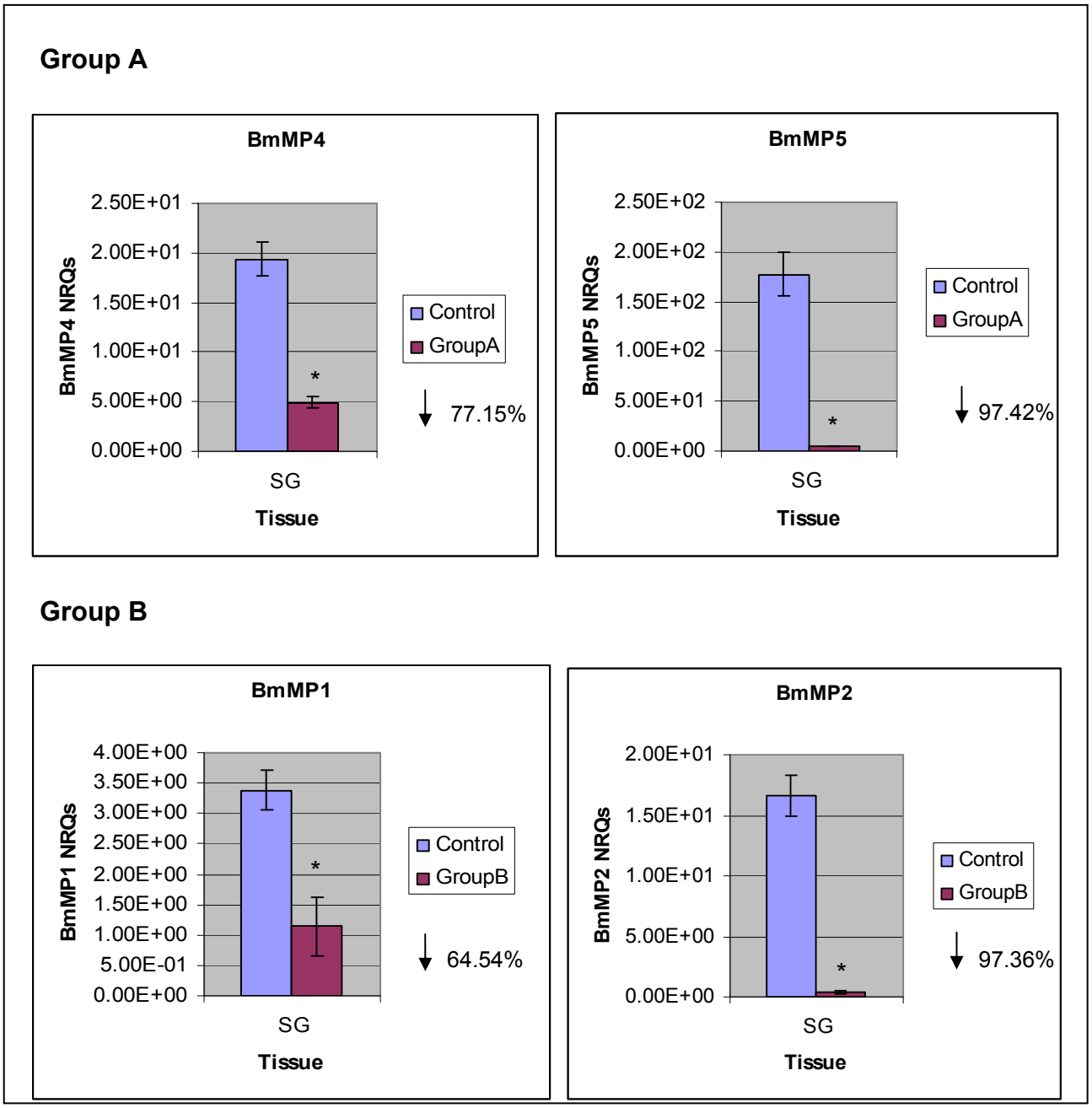

Figure 3.12. Semi-quantitative real-time PCR analysis showing the relative silenced transcript levels of BmMP4 and BmMP5 within Group A and BmMP1 and BmMP2 within Group B. Data represents mean values ± SD of the normalised relative quantities (NRQs) of 3 biological repeats, normalised to both ELF1α and PPIA. Asterisk (*) denotes the difference compared with the control group is significant as determined by the Student's *t*-test (*P< 0.01). SG; Salivary glands.

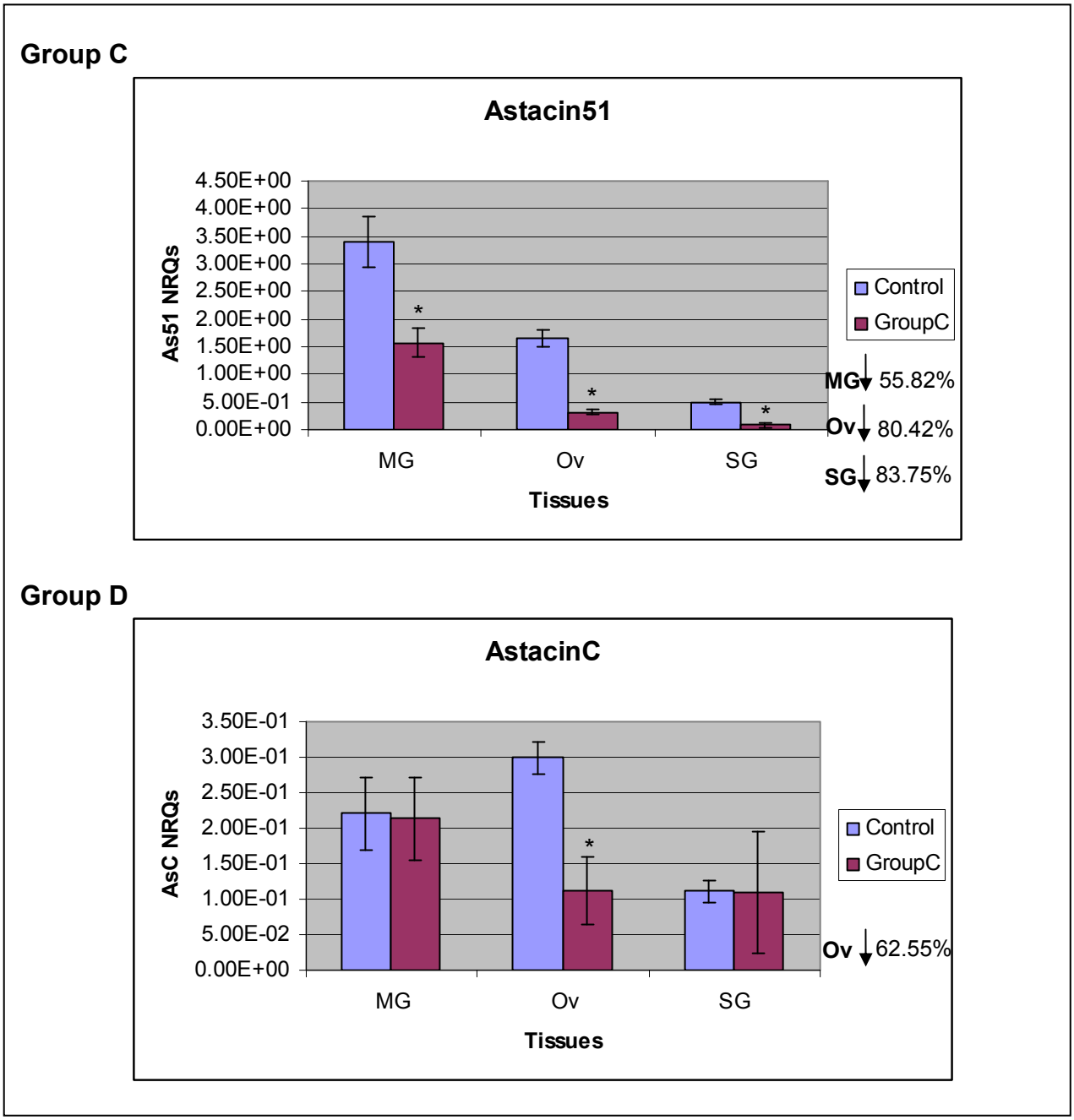

Figure 3.13. Semi-quantitative real-time PCR analysis showing the relative silenced transcript levels of As51 within Group C and AsC within Group D. Data represents mean values ± SD of the normalised relative quantities (NRQs) of 3 biological repeats, normalised to both ELF1α and PPIA. Asterisk (*) denotes the difference compared with the control group is significant as determined by the Student's *t*-test (*P< 0.01). MG;Midgut, Ov: Ovaries, SG: Salivary glands.

3.3.2.3 Non-specific gene silencing

Integrated real-time PCR reactions and data analysis revealed that significant non-specific silencing took place between As51 and AsC: injection of AsC dsRNA, in group D, not only significantly reduced the normalised levels of AsC, but also resulted in a significant reduction of As51. The normalised level of As51 was reduced with 77.79% in the midgut and 68.87% in the ovaries of Group D (Figure 3.14.). Although not found to be statistical significant, a possible *vice versa* effect cannot be excluded. Upon analysis of the alignment of the two transcripts' nucleotide sequences it is evident that there are several conserved regions. Given that initial specific RISC-target association is primarily mediated via the 5' region of the siRNA (Figure 3.3.), it has been reported that modifications in the 5' region impairs siRNA silencing, while single basepair mutations in the 3' end have been shown to have no affect on siRNA efficacy (Ameres *et al.*, 2007; Rana, 2007). Considering this, it is clear from the alignment that conserved regions (as large as 21 consecutive nucleotides) could have attributed to the non-specific silencing (Figure 3.15.).

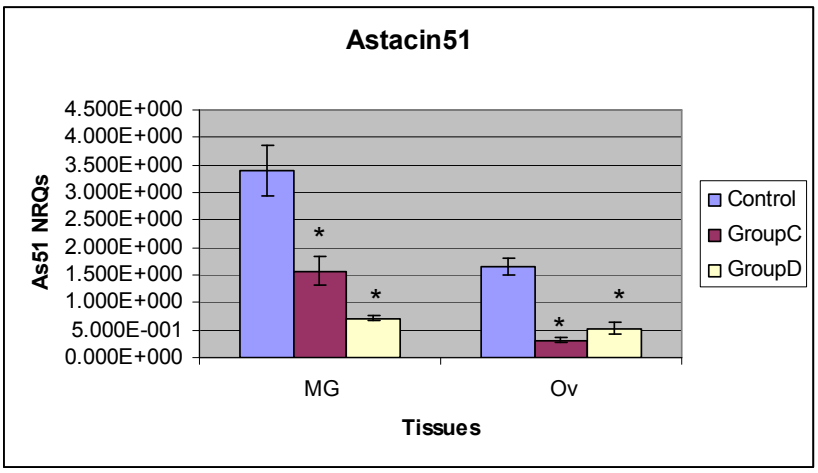

Figure 3.14. Semi-quantitative real-time PCR analysis showing the non-specific silencing of As51 in the AsC-dsRNA-injected group (Group D). Data represents mean values ± SD of the normalised relative quantities (NRQs) of 3 biological repeats, normalised to both ELF1α and PPIA. Asterisk (*) denotes the difference compared with the control group is significant as determined by the Student's *t*-test (*P< 0.01). MG: Midgut, Ov: Ovaries.

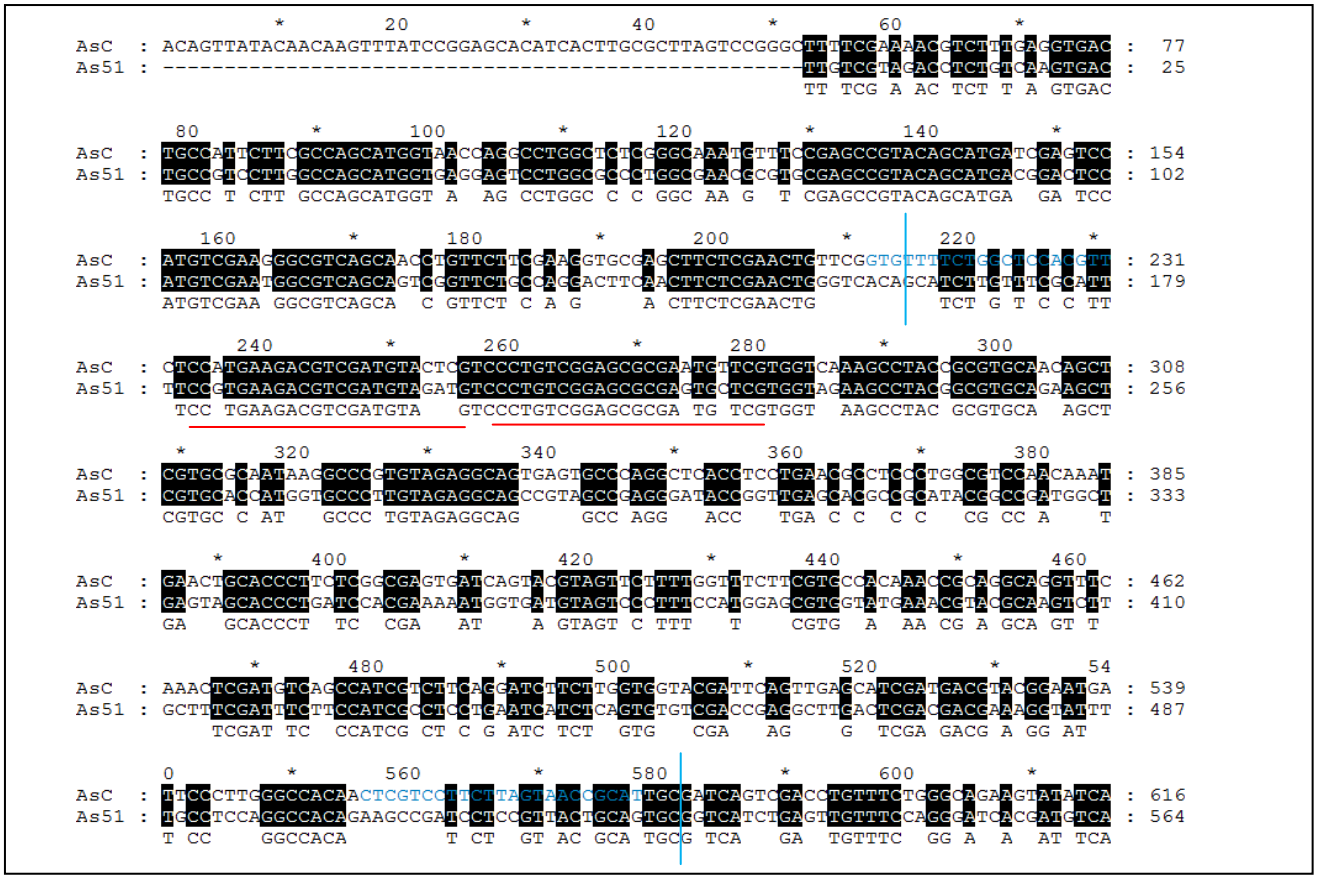

Figure 3.15. Nucleotide sequence alignment of AsC against As51. The dsRNA section of AsC is indicated between the two blue lines, with the blue regions indicating the reverse and forward primers. The red lines indicate two possible sites (up to 21 nt) which could have resulted in non-specific silencing. Black reverse background indicates conserved regions.

3.3.2.4 Differential Transcriptional Response

To investigate the differential transcriptional regulation network between the eight *R. microplus* metzincin transcripts, integrated semi-quantitative real-time PCR studies were performed. Utilising the same sets of gene-specific primers used during the silencing confirmation experiments (Table 3.11.), the expression level of each transcript was determined within all relevant tissue types of the dsRNA-injected groups. The raw (unprocessed) Cq-values were normalised relative to ELF1α and PPIA and the normalised expression level of each transcript was compared to its normalised level in the mock-injected group (Figures 3.16, 3.17 and 3.18, and Table 3.12.).

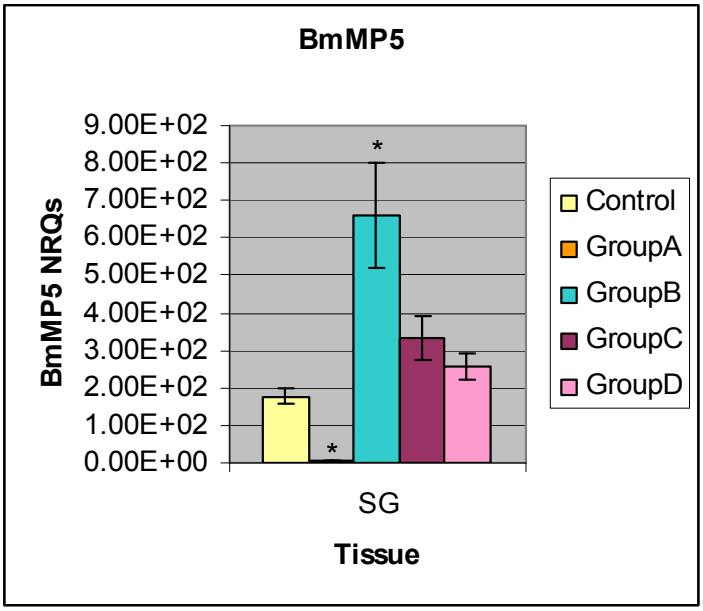

Figure 3.16. Semi-quantitative real-time PCR analysis showing the relative transcript levels of BmMP5 in the salivary glands within the control and test groups (A-D). Data represents mean values ± SD of the normalised relative quantities (NRQs) of 3 biological repeats, normalised to both ELF1α and PPIA. Asterisk (*) denotes the difference compared with the control group is significant as determined by the Student's *t*-test (*P< 0.01). SG: Salivary glands.

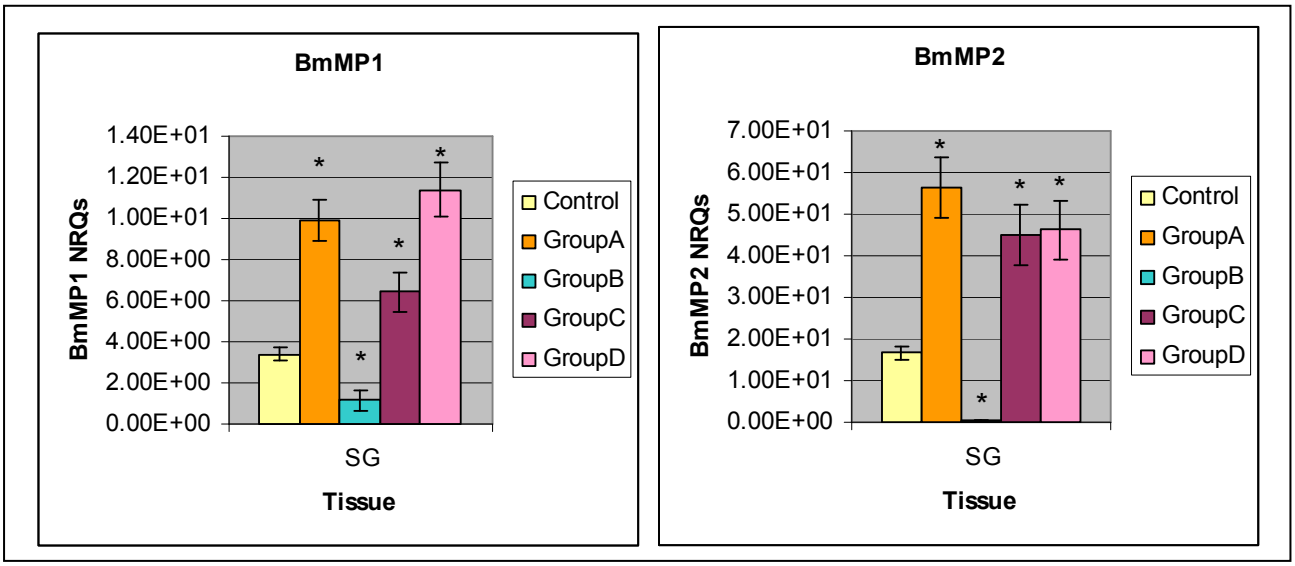

Figure 3.17. Semi-quantitative real-time PCR analysis showing the relative transcript levels of BmMP1 (left) and BmMP2 (right) in the salivary glands within the control and test groups (A-D). Data represents mean values ± SD of the normalised relative quantities (NRQs) of 3 biological repeats, normalised to both ELF1α and PPIA. Asterisk (*) denotes the difference compared with the control group is significant as determined by the Student's *t*-test (*P< 0.01). SG: Salivary glands.

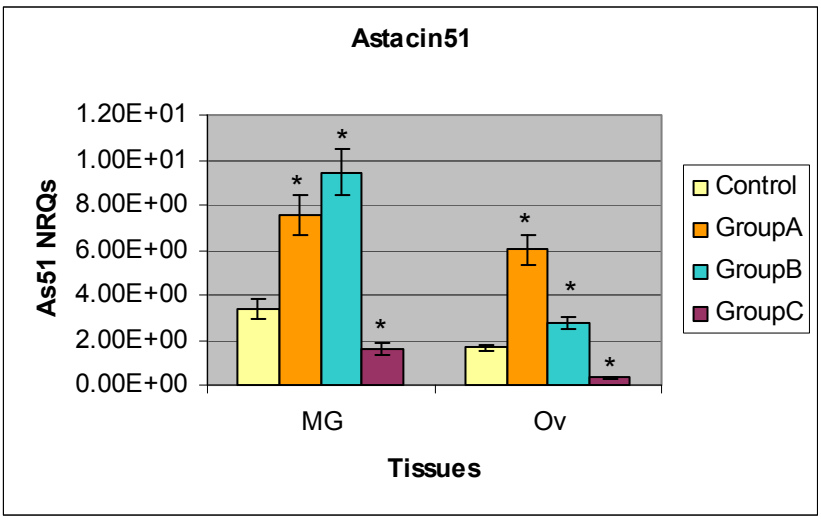

Figure 3.18. Semi-quantitative real-time PCR analysis showing the relative transcript levels of As51 in the midgut and ovaries within the control and test groups A, B and C. Data represents mean values ± SD of the normalised relative quantities (NRQs) of 3 biological repeats, normalised to both ELF1α and PPIA. Asterisk (*) denotes the difference compared with the control group is significant as determined by the Student's *t*-test (*P<0.01). MG; Midgut, Ov: Ovaries.

From these results it is evident that several fold up-regulation of the non-silenced transcripts occurred, when one/more of the other metzincin transcript(s) were silenced. With the silencing of the two salivary gland reprolysins, BmMP4 and BmMP5 (Group A), the normalised levels of the other salivary gland reprolysins, BmMP1 and BmMP2 significantly increased with 185.32% and 240.52%, (P<0.01) (Figure 3.17.), respectively. Of great interest was that the normalised level of As51 was also significantly increased with 125.44% in the midgut and 249.95% in the ovaries (P<0.01) (Figure 3.18.). The same effect was observed when the other two salivary gland reprolysins, BmMP1 and BmMP2 (Group B), were silenced. The normalised level of BmMP5 was significantly increased with 301.70% (P<0.01) (Figure 3.16.) in the salivary glands and the normalised level As51 with 188.13% in the midgut and 67.74% in the ovaries (P<0.01) (Figure 3.18.). With the silencing of both As51 (Group C) and AsC (Group D), respectively, a phenomenal cross-organ compensatory effect was observed. In this case, the normalised levels of both BmMP1 and BmMP2 were significantly increased in the salivary glands of Group C with 111.44% and 191.82% (P<0.01) and with 257.94% and 181.60% in Group D (Figure 3.17.).

		CNRQ values and percentage up or down regulation											
		BmMP1		BmMP2		BmMP4		BmMP5		As51		AsC	
Group A	SG	9.92 ±	+185.32*	56.29 ±	+240.52*	**4.95 ±**	**-77.15***	**4.89 ±**	**-97.42***	-	-	-	-
	MG	-	-	-	-	-	-	-	-	7.56 ±	+125.44*	22.05 ±	-
	O	-	-	-	-	-	-	-	-	6.01 ±	+249.95*	16.47 ±	-
Group B	SG	**1.14 ±**	**- 64.54***	**0.43 ±**	**-97.36***	-	-	661.19 ±	+301.70*	-	-	-	-
	MG	-	-	-	-	-	-	-	-	9.44 ±	+188.13*	14.78 ±	-
	O	-	-	-	-	-	-	-	-	2.73 ±	+67.74*	18.07 ±	-
Group C	SG					-	-	334.13 ±	+113.04	-	-	-	-
	MG	-	-	-	-	-	-	-	-	**1.57 ±**	**-55.82***	0.11 ±	-49.28
	O	-	-	-	-	-	-	-	-	**0.32 ±**	**-80.43***	0.29 ±	-9.34499
Group D	SG	11.38 ±	+257.94*	46.21 ±	+181.60*	-	-	256.94 ±	+47.38	-	-	-	-
	MG	-	-	-	-	-	-	-	-	0.71 ±	-77.79*	**0.21 ±**	**7.46**
	O	-	-	-	-	-	-	-	-	0.52 ±	-68.87*	**0.11±**	**-62.55***

Table 3.12. Silencing confirmation and differential transcriptional response analysis. The relative transcript levels in the salivary glands (SG), midgut (MG) and ovaries (O) of five partially fed females, 6 days after injection with dsRNA of BmMP4 and BmMP5 (Group A), BmMP1, BmMP2 and BmMP3 (Group B), As51 and As70 (Group C) and AsC (Group D) are represented by the Calibrated Normalised Relative Quantity (CNRQ)-values ±SD, normalised against ELF1-α and PPIA. The correlated percentage up (+) or down (-) regulation of each transcript within its dsRNA silenced group (in bold) and in the other groups are compared to its level in the control group. An asterisk (*) denotes if the difference is significant, as determined by the student's *t*-test (*P<0.01).

These findings suggest that a potential differential transcriptional regulation network exists between the metzincins of the same families, in example among the salivary gland reprolysins and the midgut and ovary astacins, independently. From these results it can be hypothesised that a possible counter reaction on transcriptional level, which compensates for the loss of proteolytic activity, exists.

Furthermore, these results revealed an extensive cross organ network between the reprolysins which are highly expressed in the salivary glands (BmMP1, 2, 4 and 5) and the astacins of the midgut and ovaries (As51 and AsC). Silencing of salivary gland metzincins resulted in the upregulation of the midgut and ovary metzincins and *vice versa*. Currently, the mechanism by which this network is established remains unknown and in order to unravel it requires numerous integrated transcriptome studies. In Chapter 5 different postulations regarding this differential transcriptional regulation network will be discussed.

Taking into consideration the characteristics of metzincin metalloproteases (such as their enzymatic activity and organisation into protein families) together with this hypothesis of a 'metzincin cross-organ differential transcriptional network' it can be reasoned that metzincins also come with drawbacks when considered as vaccine candidates. As proteases, they are members of a very large group of enzymes with protein redundancy and functional diversification. From these results it was evident that the metzincins were able to compensate for loss of function and thereby, most likely, prevented any severe phenotypical effect (e.g incompetence to feed and subsequent mortality). However, it will be of great interest to study the outcome when the reprolysin (salivary gland) transcripts are silenced together with the astacin (midgut and ovary) transcripts.

3.3.3 Metalloprotease activity assay

The effect of *in vivo* RNAi gene silencing of the different *R. microplus* metzincins was also investigated on protein level. Total protein was isolated from the salivary glands, midgut and ovaries from one group of five partially fed *R. microplus* females, after they were injected with injection buffer alone (Control Group), BmMP4- and BmMP5- (Group A), BmMP1-, BmMP2- and BmMP3- (Group B), As51- (Group C) and AsC- (Group D)

double stranded RNA and fed on a calf for 6 days. Due to the limited number of ticks used during the study only one biological repeat (compiled from five partially fed females) was available for the protein activity assay.

With the use of a protease fluorescent detection kit and the metalloprotease inhibitor, 1,10-phenanthroline, the total metalloprotease activity within the different tissues of each dsRNA-injected group was determined and compared to the activity in the control (mock-injected) group (Figure 3.19.). Unfortunately, the metalloprotease activity was too low in the ovaries to detect a difference between the control group and the test groups.

The metalloprotease activity assay indicated that with silencing of the reprolysin transcripts, BmMP1, BmMP2 and BmMP3 (Group B) the total metalloprotease activity decreased with 88.53% in the salivary glands and with 83.37% in the midgut. Unfortunately the salivary gland sample for Group A (BmMP4 and BmMP5) was unavailable due to the limited number of ticks used during this study. The assay, however, did indicate that with the silencing of BmMP4 and BmMP5 the total metalloprotease activity within the midgut was decreased with 57.68% compared to the control group. The activity assay also indicated that with the silencing of As51 (Group C) and AsC (Group D) the total metalloprotease activity decreased quite drastically, in comparison with control group levels. In Group C the total metalloprotease activity was decreased with 75.42% in the salivary glands and with 91.99% in the midgut, and in Group D with 53.08% and 88.24%, respectively.

At first, these results seemed to be in agreement with the gene silencing results (section 3.3.2.2): since it could have been reasoned that the silencing of the different metzincin metalloprotease genes resulted in the decreased metalloprotease activity. However, with the integrated real-time PCR studies it was shown that the different *R. microplus* metzincins, on transcript level, have the ability to compensate for one another's proteolytic loss. Thus, one would expect that the metalloprotease activity should have been almost similar to or, when considering the several fold up regulation (section 3.3.2.4), should have been even higher than the metalloprotease activity in the tissues of the control group.

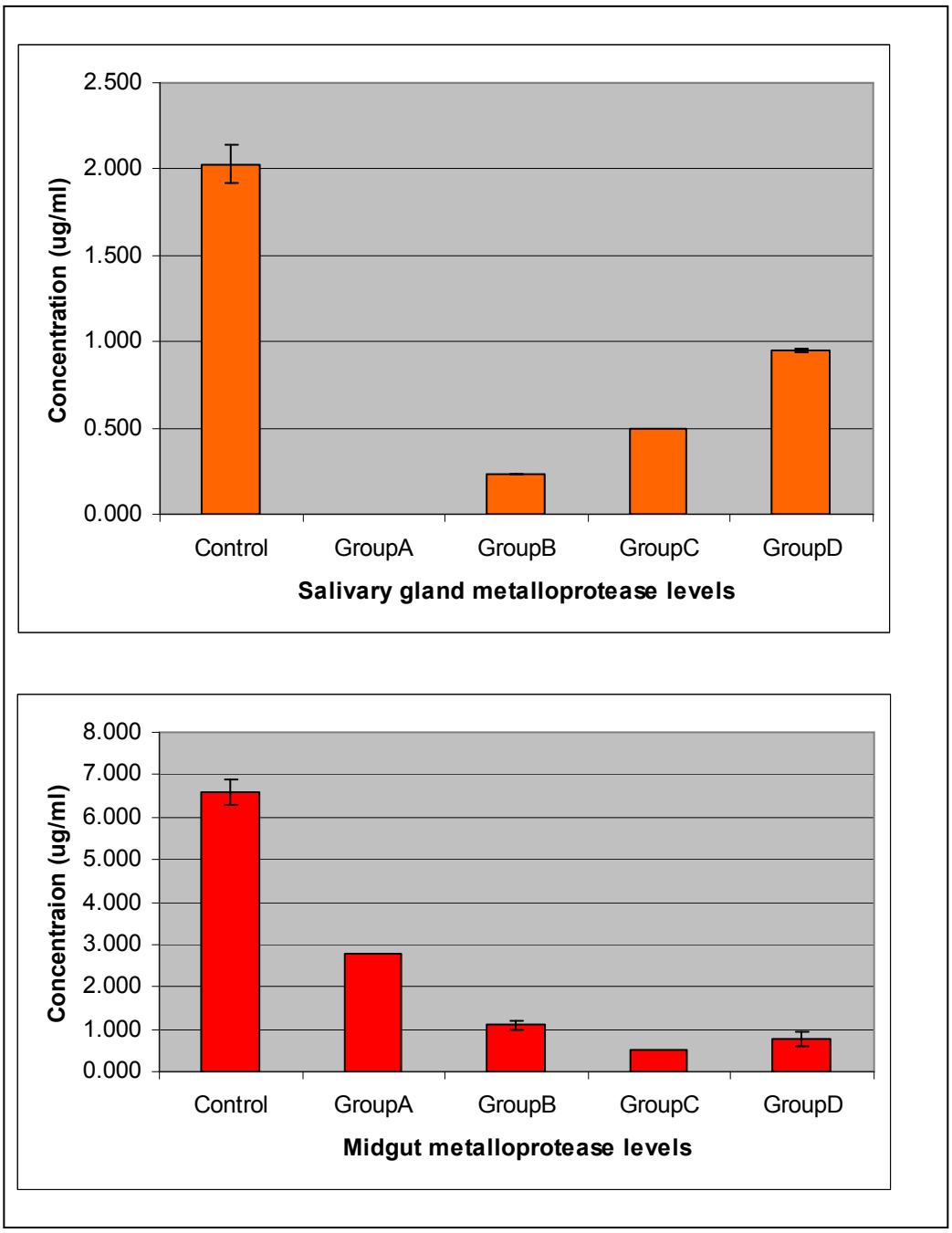

Figure 3.19. Metalloprotease activity assay analysis showing the concentration (µg/ml) of metalloproteases in the salivary gland (top) and midgut (bottom) within the control and four test groups (A-D). Data represents mean values ± SD of 2 technical repeats, from one biological repeat (which was compiled from the salivary glands and midguts of 5 partially fed female ticks).

Unfortunately, only a limited number of ticks were available for this study and as a result very limited amounts of total protein from the various dissected tissues were available. Therefore, to obtain more reliable data, the experiment must be up-scaled and repeated in its entirety. Finally, it should also be noted that although we found the 5 sequences

coding for reprolysin metzincins in the salivary glands and the 3 sequences coding for astacin metzincins in the midgut and ovaries of *R. microplus* (section 2.3.6), we cannot conclude that all metalloprotease activity reported derives solely from these metzincins. Various other *R. microplus* metalloproteases (and/or metalloprotease inhibitors) could have influenced the results. Furthermore, it has been shown that numerous arthropods, including ticks, have intact host proteins in their haemolymph after feeding (Jeffers, 2007), therefore it is also not clear if any host metalloproteases (and/or host metalloprotease inhibitors) influenced the activity results. A definitive demonstration of the enzymatic activity of each of the 8 *R. microplus* metzincins would have provided a better understanding of the effect of RNAi gene silencing and would have given great insight on how this clan of metalloproteases acts proteolytically. However, this will require a functional enzymatic/biochemical bioassay. Unfortunately, all attempts to date to perform such assays of tick metalloproteases have been unsuccessful, a general feature with metalloproteases.

3.4 Conclusion

The need to develop vaccines against tick infestations has encouraged much research for additional tick protective antigen identification. However, the difficulties, expenses and time associated with the screening of antigens and vaccination trials constitute an obstacle towards this goal. Fortunately, new and developing technologies, such as RNAi, have emerged and provide a means to overcome some of these hurdles.

RNAi has become the most widely used gene-silencing technique in ticks, where alternative approaches for genetic manipulation are not available and/or reliable. This reverse genetic tool provides a fast, less labour intensive, yet comprehensive approach to screen a large number of tick genes. In this study, *in vivo* RNAi was utilised to evaluate eight *R. microplus* metzincins (5 reprolysin and 3 astacin) as anti-*R. microplus* vital proteins for tick survival. Gene specific dsRNA was injected into freshly molted unfed females and fully engorged females to assess a possible phenotype. Gene silencing was confirmed by semi-quantitative real-time PCR in the dissected salivary glands, midguts and ovaries of the partially fed *R. microplus* females. The normalised transcript levels of the targeted genes (BmMP1, BmMP2, BmMP4, BmMP5, As51 and AsC) were reduced in the significant tissues between ~ 54 – 98% (Table 3.12.). These results are comparable with previous semi-quantitative measurements of the gene silencing effect by dsRNA-injection in *R. microplus.* In addition, these results showed that the RNAi effect was systemic, i.e. the dsRNA spread throughout the injected tick and was successfully taken up by all tissues. Recent preliminary studies indicate that the systemic RNAi effect in ticks is possibly based on a tick RdRP homologue (Kurscheid *et al.*, 2009) (Figure 3.4).

From the results it can be concluded that the silencing of the two reprolysins, BmMP1 and BmMP2, and the silencing of Astacin51, yielded the most significant phenotypes, decreasing both oviposited egg weight and oviposition efficiency. Silencing of AsC also yielded a significant phenotype, with respect to the oviposited egg mass weight. However, it must be noted that non-specific silencing occurred between AsC and As51 in Group D (Table 3.12.). Therefore it can be reasoned that the non-specific silencing of As51 in group D also contributed to the observed phenotype of Group D. With the

transovarial silencing studies the reprolysin, BmMP5 showed the most significant phenotype.

Although there are only a few published examples in which multiple antigen vaccines have been examined, impressive results have been obtained with various parasites including the tapeworm *Taenia saginata* (Lightowlers *et al.*, 1996) and one example for *R. microplus* (Willadsen *et al.*, 1996), where combinations of recombinant antigens gave more effective protection in cattle than any single antigen (Lightowlers *et al.*, 1996; Willadsen *et al.*, 1996; Willadsen, 2008). Intuitively, the situation in which the benefit of antigen cocktails should be most predictable is where multiple developmental stages or multiple physiological functions are targeted. An anti-tick vaccine from a combination of key protective antigens could reduce the rate of resistance development to any single antigen (Guerrero *et al.*, 2012b). Thus, an interesting next step would be to evaluate the phenotype when simultaneously silencing the three *R. microplus* reprolysin transcripts (BmMP1, BmMP2 and BmMP5) and the two *R. microplus* astacin transcripts (As51 and AsC), respectively. Moreover, it will be of great interest to evaluate the phenotype when simultaneously silencing the reprolysin transcripts together with the astacin transcripts. Finally, a cocktail of metzincin proteins included in a single vaccine (targeting several members of a large protein family, affecting different physiological mechanisms of the tick) should be evaluated in a vaccination trial, to establish its efficacy for the control of *R. microplus* tick infestation and ultimately for the control against different tick species and geographical strains of species.

What is of great interest, is that with the use of integrated semi-quantitative real-time PCRs and data analysis, this *in vivo* gene silencing RNAi study also provided novel insight regarding gene regulation between the different members of the metzincin clan investigated. It was found that several fold up-regulation of non-silenced transcripts occurred when other transcripts were specifically silenced (Table 3.12.). The remarkable finding was that this was a cross-organ phenomena. Upon silencing of salivary gland reprolysin metzincins, the midgut and ovary astacin metzincins were activated and *vice versa*. To fully explain this phenomena much more integrated transcriptome studies are required, which will provide great insight into *R. microplus* gene regulation during feeding and digestion. In Chapter 5 different postulations regarding this differential transcriptional regulation network will be discussed.

Finally, it can be concluded that by means of RNAi it was possible to evaluate the overall vital impact of the putative *R. microplus* reprolysin and astacin metzincins, all in order to predict if these metalloproteases can be used to produce an effective anti-tick vaccine targetting vital functions. These current results support a potential combinatorial metzincin based anti-*R. microplus* vaccine – comprised of BmMP1, BmMP2 and As51.

3.5 References

Agrawal, N., P. V. N. Dasaradhi, A. Mohmmed, P. Malhotra, R. K. Bhatnagar and S. J. Mukherjee (2003). "RNA interference: Biology, mechanism, and applications." *Microbiology and Molecular Biology Reviews* **67**(4): 657-685.

Ameres, S. L., J. Martinez and R. Schroeder (2007). "Molecular basis for target RNA recognition and cleavage by human RISC." *Cell* **130**(1): 101-112.

Aravin, A. A., G. J. Hannon and J. Brennecke (2007). "The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race." *Science* **318**(5851): 761-764.

Aung, K. M., D. Boldbaatar, R. Umemiya-Shirafuji, M. Liao, X. Xuenan, H. Suzuki, R. L. Galay, T. Tanaka and K. Fujisaki (2011). "Scavenger receptor mediates systemic RNA interference in ticks." *PloS one* **6**(12): e28407.

Barnard, A.-C., A. M. Nijhof, W. Fick, C. Stutzer and C. Maritz-Olivier (2012). "RNAi in Arthropods: Insight into the Machinery and Applications for Understanding the Pathogen-Vector Interface." *Genes* **3**: 702-741.

Bechara, G. H., M. P. Szabó, R. Z. Machado and U. F. Rocha (1988). "A technique for collecting saliva from the cattle-tick *Boophilus microplus* (Canestrini, 1887) using chemical stimulation. Environmental and temporal influences on secretion yield." *Brazilian Journal of Medical and Biological Research* **21**(3): 479-484.

Bellés, X. (2010). "Beyond *Drosophila*: RNAi *in vivo* and functional genomics in insects." *Annual Review of Entomology* **55**: 111 - 128.

Bellgard, M. I., P. M. Moolhuijzen, F. D. Guerrero, D. Schibeci, M. Rodriguez-Valle, D. G. Peterson, S. E. Dowd, R. Barrero, A. Hunter, R. J. Miller and A. E. Lew-Tabor (2012). "CattleTickBase: an integrated Internet-based bioinformatics resource for Rhipicephalus (Boophilus) microplus." *Int J Parasitol* **42**(2): 161-169.

Bernstein, E., A. A. Caudy, S. M. Hammond and G. J. Hannon (2001). "Role for a bidentate ribonuclease in the initiation step of RNA interference." *Nature* **409**(6818): 363-366.

Bowman, A. S. and J. R. Sauer (2004). "Tick salivary glands: Function, physiology and future." *Parasitology* **129**: S67-S81.

Buchon, N. and C. Vaury (2006). "RNAi: A defensive RNA-silencing against viruses and transposable elements." *Heredity* **96**: 196 - 202.

Bustin, S. A., V. Benes, J. A. Garson, J. Hellemans, J. Hugget, M. Kubista, R. Mueller, T. Nolan, M. W. Pfaffi, G. L. Shipley, J. Vandesompele and C. T. Wittwer (2009). "The MIQE guidelines: *M*inimum *I*nformation for publication of *Q*uantitative real-time PCR *E*xperiments." *Clinical Chemistry* **55**(4): 611-622.

Calixto, A., D. Chelur, I. Topalidou, X. Chen and M. Chalfie (2010). "Enhanced neuronal RNAi in *C. elegans* using SID-1." *Nature Methods* **7**(7): 554 - 559.

Carthew, R. W. and E. J. Sontheimer (2009). "Origins and mechanisms of miRNAs and siRNAs." *Cell* **136**: 642 - 655.

Caudy, A. A., R. F. Ketting, S. M. Hammond, A. M. Denli, A. M. P. Bathoorn, B. B. J. Tops, J. M. Silva, M. M. Myers, G. J. Hannon and R. H. A. Plasterk (2003). "A micrococcal nuclease homologue in RNAi effector complexes." *Nature* **425**(6956): 411-414.

Caudy, A. A., M. M. Myers, G. J. Hannon and S. M. Hammond (2002). "Fragile X-related protein and VIG associate with the RNA interference machinery." *Genes and Development* **16**(19): 2491-2496.

Cerutti, H. (2003). "RNA interference: traveling in the cell and gaining functions?" *Trends in Genetics* **19**: 39–46

Chomczynski, P. (1993). "A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples." *BioTechniques* **15**(3): 532-537.

Chung, W.-J., K. Okamura, R. Martin and E. C. Lai (2008). "Endogenous RNA interference provides a somatic defense against *Drosophila* transposons." *Current Biology* **18**(11): 795-802.

Cullen, B. R. (2006). "Is RNA interference involved in intrinsic antiviral immunity in mammals?" *Nature Immunology* **7**(6): 563-567.

de la Fuente, J., C. Almazán, U. Blas-Machado, V. Naranjo, A. J. Mangold, E. F. Blouin, C. Gortazar and K. M. Kocan (2006a). "The tick protective antigen, 4D8, is a conserved protein involved in modulation of tick blood ingestion and reproduction." *Vaccine* **24**(19): 4082-4095.

de la Fuente, J., C. Almazán, E. F. Blouin, V. Naranjo and K. M. Kocan (2005). "RNA interference screening in ticks for the identification of protective antigens." *Parasitology Research* **96**(3): 137-141.

de la Fuente, J., C. Almazán, E. F. Blouin, V. Naranjo and K. M. Kocan (2006b). "Reduction of tick infections with *Anaplasma marginale* and *A. phagocytophilum* by targeting the tick protective antigen subolesin." *Parasitology Research* **100**(1): 85-91.

de la Fuente, J., C. Almazán, V. Naranjo, E. F. Blouin and K. M. Kocan (2006c). "Synergistic effect of silencing the expression of tick protective antigens 4D8 and Rs86 in *Rhipicephalus sanguineus* by RNA interference." *Parasitology Research* **99**(2): 108-113.

de la Fuente, J. and K. M. Kocan (2006). "Strategies for the development of vaccines for control of ixodid tick species." *Parasite Immunology* **28**: 275-283.

de la Fuente, J., K. M. Kocan, C. Almazán and E. F. Blouin (2007b). "RNA interference for the study and genetic manipulation of ticks." *Trends in Parasitology* **23**(9): 427-433.

Decrem, Y., M. Mariller, K. Lahaye, V. Blasioli, J. Beaufays, K. Z. Boudjeltia, M. Vanhaeverbeek, M. Cérutti, L. Vanhamme and E. Godfroid (2007). "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**(5): 549-560.

Ding, S.-W. and O. Voinnet (2007). "Antiviral immunity directed by small RNAs." *Cell* **130**(3): 413-426.

Dlakić, M. (2006). "DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold." *Bioinformatics* **22**(22): 2711-2714.

Elbashir, S. M., W. Lendeckel and T. Tuschl (2001). "RNA interference is mediated by 21- and 22-nucleotide RNAs." *Genes and Development* **15**(2): 188-200.

Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver and C. C. Mello (1998). "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*." *Nature* **391**(6669): 806-811.

Gao, X., L. Shi, Y. Zhou, J. Cao, H. Zhang and J. Zhou (2011). "Characterization of the anticoagulant protein Rhipilin-1 from the *Rhipicephalus haemaphysaloides* tick." *Journal of Insect Physiology* **57**(2): 339 - 343.

Garcia, S., A. Billecocq, J.-M. Crance, U. Munderloh, D. Garin and M. Bouloy (2005). "Nairovirus RNA sequences expressed by a Semliki Forest virus replicon induce RNA interference in tick cells." *Journal of Virology* **79**(14): 8942-8947.

Golden, T. A., S. E. Schauer, J. D. Lang, S. Pien, A. R. Mushegian, U. Grossniklaus, D. W. Meinke and A. Ray (2002). "*Short integuments1/suspensor1/carpel factory*, a Dicer homolog, is a maternal effect gene required for embryo development in Arabidopsis." *Plant Physiology* **130**(2): 808-822.

Graham, C., P. J. Johnson and R. D. Adam (2010). *Anaerobic parasitic protozoa: Genomics and molecular biology*. United Kingdom, Caister Academic Press.

Guénin, S., M. Mauriat, J. Pelloux, O. van Wuytswinkel, C. Bellini and L. Gutierrez (2009). "Normalization of qRT-PCR data: The necessity of adopting a systematic, experimental conditions-specific, validation of references." *Journal of Experimental Botany* **60**(2): 487-493.

Guerrero, F. D., R. J. Miller and A. A. Pérez de León (2012b). "Cattle tick vaccines: Many candidate antigens, but will a commercially viable product emerge?" *International Journal for Parasitology* **42**(5): 421-427.

Haasnoot, J., E. M. Westerhout and B. Berkhout (2007). "RNA interference against viruses: Strike and counterstrike." *Nature Biotechnology* **25**(12): 1435-1443.

Haley, B. and P. D. Zamore (2004). "Kinetic analysis of the RNAi enzyme complex." *Nature Structural and Molecular Biology* **11**(7): 599-606.

Hamilton, A. J. and D. C. Baulcombe (1999). "A species of small antisense RNA in posttranscriptional gene silencing in plants." *Science* **286**(5441): 950-952.

Hammond, S. M. (2005). "Dicing and slicing: The core machinery of the RNA interference pathway." *Federation of European Biochemical Societies Letters* **579**(26): 5822-5829.

Hammond, S. M., E. Bernstein, D. Beach and G. J. Hannon (2000). "An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells." *Nature* **404**(6775): 293-296.

Hammond, S. M., S. Boettcher, A. A. Caudy, R. Kobayashi and G. J. Hannon (2001). "Argonaute2, a link between genetic and biochemical analyses of RNAi." *Science* **293**(5532): 1146-1150.

Heid, C. A., J. Stevens, K. J. Livak and P. M. Williams (1996). "Real time quantitative PCR." *Genome Research* **6**(10): 986-994.

Hellemans, J., G. Mortier, A. de Paepe, F. Speleman and J. Vandesompele (2007). "qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data." *Genome Biology* **8**(2): R19.

Hutvagner, G. and M. J. Simard (2008). "Argonaute proteins: Key players in RNA silencing." *Nature Reviews Molecular Cell Biology* **9**(1): 22-32.

Huvenne, H. and G. Smagghe (2010). "Mechanisms of dsRNA uptake in insects and potential of RNAi for pest control: A review." *Journal of Insect Physiology* **56**(3): 227-235.

Jaskiewicz, L. and W. Filipowicz (2008). "Role of Dicer in posttranscriptional RNA silencing." *Current Topics in Microbiology and Immunology* **320**: 77-97.

Jose, A. M., J. J. Smith and C. P. Hunter (2009). "Export of RNA silencing from *C. elegans* tissues does not require the RNA channel SID-1." *Proceedings of the National Academy of Sciences of the United States of America* **106**(7): 2283 - 2288.

Karim, S., V. G. Ramakrishnan, J. S. Tucker, R. C. Essenberg and J. R. Sauer (2004). "*Amblyomma americanum* salivary glands: Double-stranded RNA-mediated gene silencing of synaptobrevin homologue and inhibition of PGE2 stimulated protein secretion." *Insect Biochemistry and Molecular Biology* **34**(4): 407-413.

Karim, S., E. Troiano and T. N. Mather (2010). "Functional genomics tool: Gene silencing in *Ixodes scapularis* eggs and nymphs by electroporated dsRNA." *BMC Biotechnology* **10**(1).

Kawamata, T. and Y. Tomari (2010). "Making RISC." *Trends in Biochemical Sciences* **35**(7): 368-376.

Khvorova, A., A. Reynolds and S. D. Jayasena (2003). "Functional siRNAs and miRNAs exhibit strand bias." *Cell* **115**(2): 209-216.

Knight, S. W. and B. L. Bass (2001). "A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*." *Science* **293**(5538): 2269-2271.

Kocan, K. M., R. Manzano-Roman and J. de la Fuente (2007). "Transovarial silencing of the subolesin gene in three-host ixodid tick species after injection of replete females with subolesin dsRNA." *Parasitology Research* **100**(6): 1411-1415.

Kocan, K. M., J. Yoshioka, D. E. Sonenshine, J. de la Fuente, S. M. Ceraul, E. F. Blouin and C. Almazán (2005). "Capillary tube feeding system for studying tick-pathogen interactions of *Dermacentor variabilis* (Acari: Ixodidae) and *Anaplasma marginale* (Rickettsiales: Anaplasmataceae)." *Journal of Medical Entomology* **42**(5): 864-874.

Kubista, M., J. M. Andrade, M. Bengtsson, A. Forootan, J. Jonák, K. Lind, R. Sindelka, R. Sjöback, B. Sjögreen, L. Strömbom, A. Ståhlberg and N. Zoric (2006). "The real-time polymerase chain reaction." *Molecular Aspects of Medicine* **27**(2-3): 95-125.

Kurscheid, S., A. E. Lew-Tabor, M. Rodriguez Valle, A. G. Bruyeres, V. J. Doogan, U. G. Munderloh, F. D. Guerrero, R. A. Barrero and M. I. Bellgard (2009). "Evidence of a tick RNAi pathway by comparative genomics and reverse genetics screen of targets with known loss-of-function phenotypes in *Drosophila*." *BMC Molecular Biology* **10**(26).

Lau, P.-W., C. S. Potter, B. Carragher and I. J. MacRae (2009). "Structure of the human dicer-TRBP complex by electron microscopy." *Structure* **17**(10): 1326-1332.

Lee, Y., I. Hur, S.-Y. Park, Y.-K. Kim, R. S. Mi and V. N. Kim (2006). "The role of PACT in the RNA silencing pathway." *EMBO Journal* **25**(3): 522-532.

Lightowlers, M. W., R. Rolfe and C. G. Gauci (1996). "*Taenia saginata*: Vaccination against cysticercosis in cattle with recombinant oncosphere antigens." *Experimental Parasitology* **84**(3): 330-338.

Lingel, A., B. Simon, E. Izaurralde and M. Sattler (2004). "Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain." *Nature Structural and Molecular Biology* **11**(6): 576-577.

Liu, Q. and Z. Paroo (2010). "Biochemical principles of small RNA pathways." *Annual Review of Biochemistry* **79**: 295-319.

Ma, E., I. J. MacRae, J. F. Kirsch and J. A. Doudna (2008). "Autoinhibition of human Dicer by its internal helicase domain." *Journal of Molecular Biology* **380**(1): 237-243.

Ma, J.-B., K. Ye and D. J. Patel (2004). "Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain." *Nature* **429**(6989): 318-322.

Ma, J.-B., Y.-R. Yuan, G. Meister, Y. Pei, T. Tuschl and D. J. Patel (2005). "Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein." *Nature* **434**(7033): 666-670.

MacRae, I. J., K. Zhou, F. Li, A. Repic, A. N. Brooks, W. Z. Cande, P. D. Adams and J. A. Doudna (2006). "Structural basis for double-stranded RNA processing by Dicer." *Science* **311**(5758): 195-198.

Matranga, C., Y. Tomari, C. Shin, D. P. Bartel and P. D. Zamore (2005). "Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes." *Cell* **123**(4): 607-620.

Matranga, C. and P. D. Zamore (2007). "Small silencing RNAs." *Current Biology* **17**(18): R789-R793.

McManus, M. T. and P. A. Sharp (2002). "Gene silencing in mammals by small interfering RNAs." *Nature Reviews Genetics* **3**(10): 737-747.

Moolhuijzen, P. M., A. E. Lew-Tabor, J. A. Morgan, M. R. Valle, D. G. Peterson, S. E. Dowd, F. D. Guerrero, M. I. Bellgard and R. Appels (2011). "The complexity of Rhipicephalus (Boophilus) microplus genome characterised through detailed analysis of two BAC clones." *BMC research notes* **4**: 254.

Napoli, C., C. Lemieux and R. Jorgensen (1990). "Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans." *Plant Cell* **2**(4): 279-289.

Nijhof, A. M., A. Taoufik, J. de la Fuente, K. M. Kocan, E. de Vries and F. Jongejan (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**(6): 653-662.

Nykänen, A., B. Haley and P. D. Zamore (2001). "ATP requirements and small interfering RNA structure in the RNA interference pathway." *Cell* **107**(3): 309-321.

Obbard, D. J., K. H. J. Gordon, A. H. Buck and F. M. Jiggins (2009). "The evolution of RNAi as a defence against viruses and transposable elements." *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**(1513): 99-115.

Pal, U., X. Li, T. Wang, R. R. Montgomery, N. Ramamoorthi, A. M. de Silva, F. Bao, X. Yang, M. Pypaert, D. Pradhan, F. S. Kantor, S. Telford, J. F. Anderson and E. Fikrig (2004). "TROSPA, an *Ixodes scapularis* receptor for *Borrelia burgdorferi*." *Cell* **119**(4): 457-468.

Parker, J. S., S. M. Roe and D. Barford (2004). "Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity." *EMBO Journal* **23**(24): 4727-4737.

Pham, J. W., J. L. Pellino, Y. S. Lee, R. W. Carthew and E. J. Sontheimer (2004). "A Dicer-2-dependent 80S complex cleaves targeted mRNAs during RNAi in Drosophila." *Cell* **117**(1): 83-94.

Provost, P., D. Dishart, J. Doucet, D. Frendewey, B. Samuelsson and O. Rådmark (2002). "Ribonuclease activity and RNA binding of recombinant human Dicer." *EMBO Journal* **21**(21): 5864-5874.

Ramakrishnan, V. G., M. N. Aljamali, J. R. Sauer and R. C. Essenberg (2005). "Application of RNA interference in tick salivary gland research." *Journal of Biomolecular Techniques* **16**(4): 297-305.

Ramamoorthi, N., S. Narasimhan, U. Pal, F. Bao, X. F. Yang, D. Fish, J. Anguita, M. V. Norgard, F. S. Kantor, J. F. Anderson, R. A. Koski and E. Fikrig (2005). "The Lyme disease agent exploits a tick protein to infect the mammalian host." *Nature* **436**(573-577).

Rana, T. M. (2007). "Illuminating the science: Understanding the structure and function of small RNAs." *Nature Reviews Molecular Cell Biology* **8**(1): 23-36.

Saito, K. C., G. H. Bechara, É. T. Nunes, P. R. de Oliveira, S. E. Denardi and M. I. C. Mathias (2005). "Morphological, histological, and ultrastructural studies of the ovary of the cattle-tick *Boophilus microplus* (Canestrini, 1887) (Acari: Ixodidae)." *Veterinary Parasitology* **129**(3-4): 299-311.

Saleh, M.-C., R. P. van Rij, A. Hekele, A. Gillis, E. Foley, P. H. O'Farrell and R. Andino (2006). "The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing." *Nature Cell Biology* **8**(8): 793-802.

Sasaki, T. and N. Shimizu (2007). "Evolutionary conservation of a unique amino acid sequence in human DICER protein essential for binding to Argonaute family proteins." *Gene* **396**(2): 312-320.

Sasaki, T., A. Shiohama, S. Minoshima and N. Shimizu (2003). "Identification of eight members of the Argonaute family in the human genome." *Genomics* **82**(323-330).

Sashital, D. G. and J. A. Doudna (2010). "Structural insights into RNA interference." *Current Opinion in Structural Biology* **20**(1): 90-97.

Schwarz, D. S., G. Hutvánger, T. Du, Z. Xu, N. Aronin and P. D. Zamore (2003). "Asymmetry in the assembly of the RNAi enzyme complex." *Cell* **115**(2): 199-208.

Shimell, M. J., E. L. Ferguson, S. R. Childs and M. B. O'Connor (1991). "The Drosophila dorsal-ventral patterning gene *tolloid* is related to human bone morphogenetic protein 1." *Cell* **67**(3): 469-481.

Sijen, T., and Plasterk, R. H. A. (2003). "Transposon silencing inthe Caenorhabditis elegans germ line by natural RNAi." *Nature* **426**: 310–314.

Sijen, T. and R. H. A. Plasterk (2003). "Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi." *Nature* **426**(6964): 310-314.

Soares, C. A. G., C. M. R. Lima, M. C. Dolan, J. Piesman, C. B. Beard and N. S. Zeidner (2005). "Capillary feeding of specific dsRNA induces silencing of the isac gene in nymphal *Ixodes scapularis* ticks." *Insect Molecular Biology* **14**(4): 443-452.

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

Song, J.-J., J. Liu, N. H. Tolia, J. Schneiderman, S. K. Smith, R. A. Martienssen, G. J. Hannon and L. Joshua-Tor (2003). "The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes." *Nature Structural Biology* **10**(12): 1026-1032.

Song, J.-J., S. K. Smith, G. J. Hannon and L. Joshua-Tor (2004). "Crystal structure of argonuate and its implications for RISC slicer activity." *Science* **305**(5689): 1434-1437.

Stöcker, W., F.-X. Gomis-Rüth, W. Bode and R. Zwilling (1993). "Implications of the three-dimensional structure of astacin for the structure and function of the astacin family of zinc-endopeptidases." *European Journal of Biochemistry* **214**(1): 215-231.

Tijsterman, M., R. C. May, F. Simmer, K. L. Okihara and R. H. A. Plasterk (2004). "Genes required for systemic RNA interference in *Caenorhabditis elegans*." *Current Biology* **14**(2): 111-116.

Tomari, Y., T. Du, B. Haley, D. S. Schwarz, R. Bennett, H. A. Cook, B. S. Koppetsch, W. E. Theurkauf and P. D. Zamore (2004a). "RISC assembly defects in the *Drosophila* RNAi mutant *armitage*." *Cell* **116**(6): 831-841.

Tomari, Y., C. Matranga, B. Haley, N. Martinez and P. D. Zamore (2004b). "A protein sensor for siRNA asymmetry." *Science* **306**(5700): 1377-1380.

Tomari, Y. and P. D. Zamore (2005). "Perspective: Machines for RNAi." *Genes and Development* **19**(5): 517-529.

Tuschl, T., P. D. Zamore, R. Lehmann, D. P. Bartel and P. A. Sharp (1999). "Targeted mRNA degradation by double-stranded RNA *in vitro*." *Genes and Development* **13**(24): 3191-3197.

Twining, S. S. (1984). "Fluorescein Isothiocyanate-labeled casein assay for proteolytic enzymes." *Analytical Biochemistry* **143**(1): 30-34.

van den Berg, A., J. Mols and J. Han (2008). "RISC-target interaction: Cleavage and translational suppression." *Biochimica et Biophysica Acta* **1779**(11): 668-677.

Wang, H.-W., C. Noland, B. Siridechadilok, D. W. Taylor, E. Ma, K. Felderer, J. A. Doudna and E. Nogales (2009). "Structural insights into RNA processing by the human RISC-loading complex." *Nature Structural and Molecular Biology* **16**(11): 1148-1153.

Wang, M., F. D. Guerrero, G. Pertea and V. M. Nene (2007). "Global comparative analysis of ESTs from the southern cattle tick, *Rhipicephalus* (*Boophilus*) *microplus*." *BMC Genomics* **8**: 368.

Willadsen, P. (2008). "Antigen cocktails: Valid hypothesis or unsubstantiated hope?" *Trends in Parasitology* **24**(4): 164-167.

Willadsen, P., D. Smith, G. Cobon and R. V. McKenna (1996). "Comparative vaccination of cattle against *Boophilus microplus* with recombinant antigen Bm86 alone or in combination with recombinant Bm91." *Parasite Immunology* **18**(5): 241-246.

Winston, W. M., C. Molodowitch and C. P. Hunter (2002). "Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1." *Science* **295**(5564): 2456-2459.

Yan, K. S., S. Yan, A. Farooq, A. Han, L. Zeng and M.-M. Zhou (2003). "Structure and conserved RNA binding of the PAZ domain." *Nature* **426**(6965): 469-474.

Yigit, E., P. J. Batista, Y. Bei, K. M. Pang, C.-C. G. Chen, N. H. Tolia, L. Joshua-Tor, S. Mitani, M. J. Simard and C. C. Mello (2006). "Analysis of the *C. elegans* argonaute family reveals that distinct argonautes act sequentially during RNAi." *Cell* **127**(4): 747-757.

Zamore, P. D. and B. Haley (2005). "Ribo-gnome: The big world of small RNAs." *Science* **309**(5740): 1519-1524.

Zamore, P. D., T. Tuschl, P. A. Sharp and D. P. Bartel (2000). "RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals." *Cell* **101**(1): 25-33.

Zhang, H., F. A. Kolb, V. Brondani, E. Billy and W. Filipowicz (2002). "Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP." *EMBO Journal* **21**(21): 5875-5885.

Zhang, H., F. A. Kolb, L. Jaskiewicz, E. Westhof and W. Filipowicz (2004). "Single processing center models for human Dicer and bacterial RNase III." *Cell* **118**(1): 57-68.

Zhou, J., M. Liao, T. Hatta, M. Tanaka, X. Xuan and K. Fujisaki (2006b). "Identification of a follistatin-related protein from the tick *Haemaphysalis longicornis* and its effect on tick oviposition." *Gene* **372**(1-2): 191-198.

Chapter 4

Recombinant protein expression for vaccination

4.1 Introduction

The widespread resistance of *R. microplus* to expensive and contaminating acaricides has forced investigations to more cost effective and efficient control strategies. One strategy to address the enduring *R. microplus* problem is to identify, characterise and evaluate new tick proteins that can serve as protective antigens in vaccines. Recent advances in technologies, such as next-generation sequencing, provide immense throughputs and yields of data on parasitic vectors and pathogens. Various algorithms and bio-informatic tools have been developed and are available to 'mine' data for putative vaccine candidates or targets, by characterising and predicting specific properties of antigenic proteins such as immunogenic epitopes. Reverse genetic technologies, such RNAi again allows for preliminary evaluation of an antigen. By means of gene specific silencing and phenotype assessment, RNAi provides insight into the physiological importance of a target in the parasite vector or pathogen. However, the target's protective capability can only be assertively verified by challenging the host and assessing the protective ability of an antigen. This however, requires large amounts of the protein of interest. Most proteins are not abundantly expressed in the natural host and therefore cannot be isolated in sufficient quantities for downstream studies. Heterologous expression (expression of the foreign gene of interest in an expression system) has bridge this gap and is of vital importance in molecular science.

Protein expression systems that have been developed for heterologous recombinant expression include prokaryotic (bacterial e.g. *E. coli*), eukaryotic (yeast, plants, mammalian cells and insect cells/baculovirus) and cell-free systems. Each of these systems has its strengths and weaknesses concerning cost, production time, ease of use, yield, contamination risk, proper protein folding and post-translational modifications (Figure 4.1). Since there is no universal procedure for heterologous expression, the type of protein (soluble or membrane bound) and its experimental purpose (structural study, activity assay or preparation of recombinant vaccine) should be assessed to establish the choice of expression system. *E. coli* is usually the first choice, based on its

ease of being genetically manipulated and the availability of a variety of suitable cloning vectors and host strains.

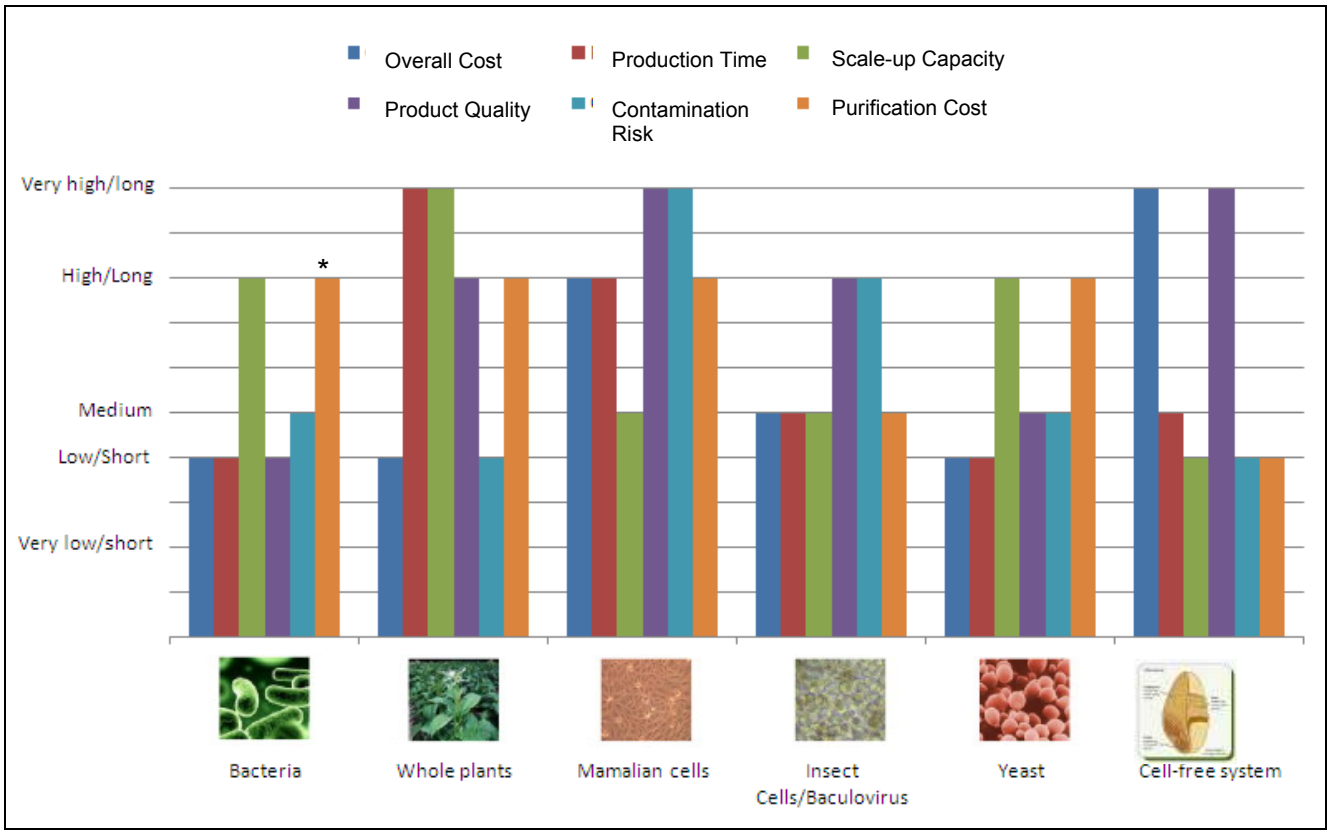

Figure 4.1. Comparison of key production parameters of the different expression systems used for recombinant proteins. (Adapted from Xu *et al.*, 2011) Asterisk (*) indicates the verity that the purification cost of the bacteria cell culture system, such as MSP1α will become low if whole membrane fraction, containing recombinant protein is utilised for subsequent studies.

4.1.1. Protein expression systems used for recombinant protein production

4.1.1.1 *Escherichia coli*

Bacterial expression is by far the most preferred and extensively used host for the production of heterologous recombinant protein. The vast utilisation of *E. coli* is mainly based on its low cost, ease to grow cultures to high cell densities (over 100 g/L) and simple nutritional requirements that can be achieved with fully defined simple media (Palomares *et al.*, 2002). An immense amount of knowledge regarding this prokaryotic organism's molecular biology, physiology and biochemistry is available (Baneyx, 1999;

Tolmasky *et al.*, 1999), thereby allowing controlled genetic manipulation. *E. coli* expression can be optimised to ensure the stability of the plasmid vector, the absence of destructive natural proteases and availability of the relevant genetic elements, such as DE3 in BL21 cells which allow the induction of T7 RNA polymerase expression (Sørensen and Mortensen, 2005).

Further advantages of *E. coli* as expression system include its big range of available strains and compatible expression vectors (e.g. pET or pBAD) with multiple fusion tags (e.g. glutathione S-transferase or hexahistidine) which is under control of different promoters (e.g. T7 or trc). However, there are also limitations to using *E. coli* as an expression host. *E. coli* is incapable of producing post-translation modification, which naturally takes place in eukaryotic cells, such as glycosylation that can be critical for the production of correctly folded active protein. Another major limitation of this system is the expression of proteins as insoluble protein, insulated in inclusion bodies, as a result of metabolic stress (Carrió and Villaverde, 2002).

Low recombinant protein production in *E. coli* can be addressed by several counter strategies, mediating solutions to obstacles such as codon bias or post-transcriptional modifications (such as disulphide bond formation and protein phosphorylation). To improve the solubility of recombinant protein, N- or/ and C-terminal fusion tags such as Glutamine-S-transferase (GST), hexahistidine and various others (Esposito and Chatterjee, 2006) can be employed. Moreover, to promote expression of soluble recombinant proteins, various expression conditions, such as lower incubation temperatures and different media composition can be explored (Sahdev *et al.*, 2008). Codons that occur in high frequency in the organism from which the foreign gene originate, but in low frequencies in the *E.coli* host will often result in a condition whereby the pool of tRNA for that codon will become depleted. Depletion of rare tRNAs during translation of foreign mRNA is one of the main reasons for the reduction in growth of the host cells and can easily cause an overall toxic response (de Boer and Kastelein, 1986; Akashi, 2001). Gene expression is also controlled by the rate at which mRNA is degraded, thus *E. coli* host strains that are deficient in specific RNases consequently improve expression by limiting mRNA degradation (Sørensen and Mortensen, 2005).

4.1.1.2 Yeast

Yeast expression systems offer a very powerful alternative to other expression systems. Its great potential lies in the fact that it combines the advantages of unicellular organisms (e.g. ease of genetic manipulation and growth) with the protein processing capability of a eukaryotic host (e.g. post-translational modification, protein folding and assembly), together with the absence of endotoxins.

Yeasts that have been used to express eukaryotic genes include *Saccharomyces cerevisiae, Pichia pastoris, Schizosaccharomyces pombe, Hansela polymorpha, Kluyveromyces lactis* and *Yarrowia lipolytica*. Of these, the first two are the most widely used. *S. cerevisiae* was the first of the yeast to be used as host, as its molecular biology and physiology is the best characterised (Domînquez *et al.*, 1998). With carefully controlled growth fermenters, impressive secreted protein yields of 10-75 mg/L have been reported utilising *S. cerevisiae* as expression host (Miles *et al.*, 2002). However, the fact that *S. cerevisiae* has primitive glycosylation and amino acid modification systems, resulting in hyperglycosylation and inaccurate O- and N-linked glycosylation of the heterologous proteins, has resulted in this yeast system not always being an optimal host for recombinant protein expression (Tanner and Lehle, 1987; Porro and Mattanovich, 2004). The methylotrophic budding yeast *P. pastoris* is in general the most popular yeast host of choice for the expression of recombinant vaccines and drug targets of disease causing organisms, including *Plasmodium falciparum* and various tick species (García-García *et al.*, 1998; Zou *et al.*, 2003). As for *S. cerevisiae*, *P. pastoris* is driven by an inducible promoter, the methanol-inducible alcohol oxidase promoter AOX1. Together with several co-factors, the redox environment of *P. pastoris* not only permits post-translational modifications but also allow high yields of secreted recombinant protein into the cell culture medium (up to 1g/L), circumventing the formation of inclusion bodies that in turn allows for purification from less complex fraction (Cereghino *et al.*, 2002). These characteristic allow for the expression of vaccine candidate proteins that rely critically on conformational epitopes to generate a protective immune response.

Nevertheless, there are limitations to *P. pastoris* expression host. Disulphide bonding within proteins may be heterogenous, producing several structural conformers of one

protein, requiring refolding to prevent incorrect conformations of epitopes (Birkholtz *et al.*, 2008). Strong expression promoters can lead to the overproduction of protein leading to depletion of precursors and nutritional compounds needed for energy (Balamurugan *et al.*, 2007). Furthermore, cells with over production of protein can also lead to the aggregation of protein in the endoplasmic reticulum, which finally shut down the secretory pathway (Cereghino and Cregg, 1999). Altogether this may lead to difficult, time-consuming and ultimately expensive optimization.

4.1.1.3 Baculovirus-insect cell systems

Since Summer and Smith demonstrated in the early 1980s that the double stranded DNA genomes of baculoviruses could be manipulated to encode foreign proteins, baculovirus-insect cell systems have become a popular tool for the production of complex proteins (Kost *et al.*, 2005). With numerous baculovirus systems and a variety of transfer vectors available, this technology holds great potential, which has already been illustrated by successful application in commercial manufacturing of various veterinary and human vaccines (Mena and Kamen, 2011; Cox, 2012). The key advantage of this recombinant protein expression platform is that a universal “plug and play” process may be used for producing a broad range of drug targets and recombinant vaccines, while offering the potential for low manufacturing costs (Mena and Kamen, 2011).

As the baculovirus expression system utilises eukaryotic insect cells, it has the capability and advantages of eukaryotic expression systems to produce high yields of secreted protein, due to its ability to incorporate post-translational modifications such as glycosylation (James *et al.*, 1995), phosphorylation (Hericourt *et al.*, 2000) and disulphide bond formation (Hodder *et al.*, 1996). This system is also able to recognise eukaryotic targeting signals which drive expression and processing of diverse classes of proteins, including secreted, membrane or cytoplasmic proteins (Luckow and Summers, 1988). The eukaryotic potential allows for correct folding and assembly of recombinant proteins, thereby displaying immunogenic regions in its accurate state to be recognised by conformational specific antibodies. Expression of hetro-oligomeric protein complexes can also be established by simultaneously infecting cells with two or more recombinant

baculoviruses or by infecting cells with recombinant virus containing two or more expression cassettes (Kakker *et al.*, 1999).

Baculoviruses are a family of large rod-shaped viruses that have a narrow host range, restricted to a limited number of closely related insect species (including lepidopteran, dipteran and hymenopteran) (Tinsley and Harrap, 1978). Within these insects, baculoviruses cause fatal diseases, which initially lead to the exploitation of applying these viruses as bio-control agents (Cox, 2012). Although these viruses are capable of entering vertebrate cells, they are incapable of replicating in them, making them a safer alternative to conventional virus-vectors and accelerating the approval for commercial use (Murphy *et al.*, 2004). The most commonly used baculovirus is the lytic *Autographa californica* nuclear polyhedrosis virus. Under the transcriptional control of a strong polyhedrin promoter, the virus abundantly expresses heterologous genes during the late stages of infection. *A. californica* host range include the most commonly used insect cell culture systems *Spodoptera frugiperda* and *Trichoplusia ni* derived from the fall armyworm and cabbage looper, respectively (Tinsley and Harrap, 1978). Though *A. californica* replicate well within these cells, manipulation of the virus genome is complex due to its large genome size (134 kbp), rendering direct insertion of foreign genes. Thus, the main disadvantage of the baculovirus-insect cell system is that a tedious, technically challenging and time consuming process is required to generate recombinant baculovirus (Trowitzsch *et al.*, 2010). Different strategies include lengthy homologous recombination within the insect cells or a more generally used two step process involving: a) cloning of the gene of interest into a donor plasmid for insertion into the virsus genome and b) subsequent isolation of recombinant virus from non-recombinant virus. The most popular used method and the basis of several other methods was first developed in 1993 by Luckow and colleagues (1993) and marketed as Bac-to-Bac™ by Invitrogen (See Figure 4. 2 for a full description). Briefly, it is based on site-specific transposition of an expression cassette (from a donor plasmid) into a baculovirus shuttle vector (bacmid), at the polyhedron locus, propagated in competent bacterial cells. With appropriate antibiotics recombinant clones are selected, DNA is extracted from selected positive clones, and is used to transfect insect cells for expression (Luckow *et al.*, 1993).

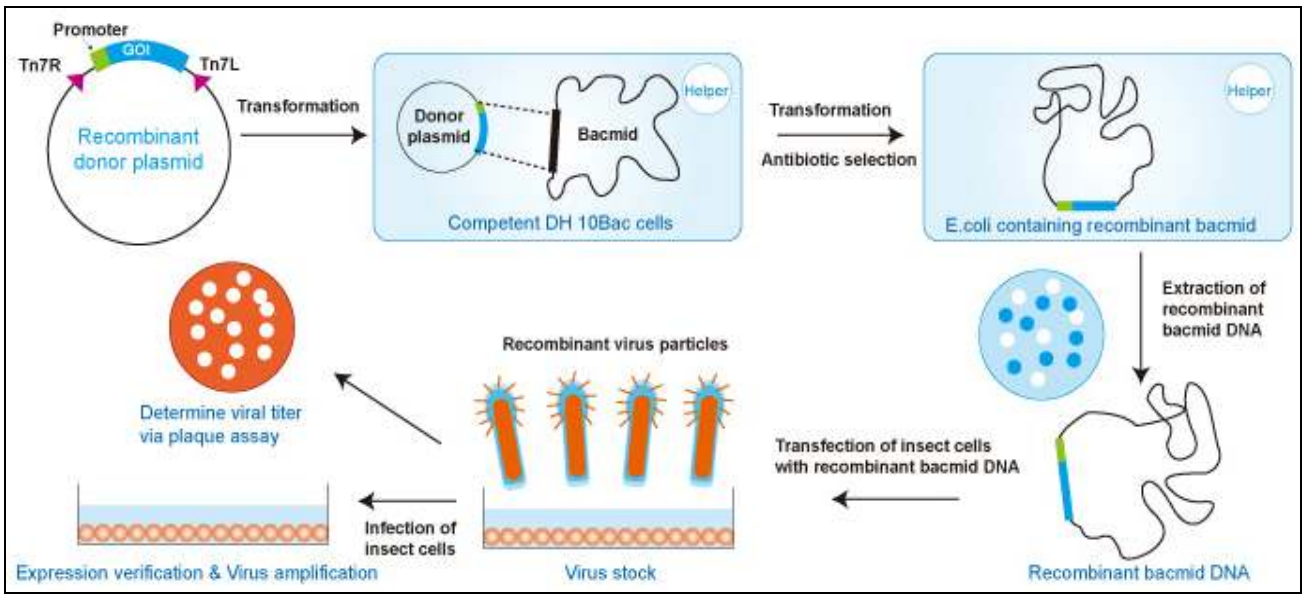

Figure 4.2. Schematic representation of the generation of recombinant baculoviruses and gene expression with the Bac-to-Bac® Expression System. The gene of interest is cloned into a donor plasmid and the recombinant plasmid is transformed into DH10Bac™ competent cells, which contain the bacmid with a mini-*att*Tn7 target site and the helper plasmid. The mini-Tn7 element on the donor plasmid can transpose to the mini-*att*Tn7 target site on the bacmid in the presence of transposition proteins provided by the helper plasmid. Colonies containing recombinant bacmids are identified by disruption of the *lac*Zα gene. High molecular weight mini-prep DNA is prepared from selected *E. coli* clones containing the recombinant bacmid and this DNA is then used to transfect insect cells.

Apart from its main technical drawback, the lack of a straight forward easy-to-use procedure to generate recombinant virus, the baculovirus-insect cell expression system has a few additional limitations regarding the expression of some eukaryotic target proteins. These include the inability of the system to produce recombinant GPI-anchor proteins and in some cases incomplete post-translational processing, as expression occurs mainly during the late stages of the viral lifecycle (Gowda and Davidson, 1999). Nevertheless, the baculovirus-insect cell expression system is a valuable technology and together with the potential to be up-scaled and automated, it has broad implications for pathogen control and disease treatment.

4.1.1.4 Mammalian cells

Although initially hurdles such as the cellular fragility and the complex nutritional requirements of mammalian cells had to be overcome, mammalian cell cultures are currently extensively exploited for recombinant protein production (Ramírez *et al.*,

1990). The main benefit of mammalian cell expression systems is that a cellular environment closer to the physiological one is provided.

Mammalian cells are well characterised on molecular level and have a versatile availability of protocols and reagents for heterologous gene expression. Expression of proteins in mammalian cell systems holds several advantages over other systems, as codon optimisation is rarely needed and because it has the capability to execute all post-transcriptional modifications necessitated for proper protein folding and correct sub-cellular targeting (including GPI-anchored proteins) (Hunt, 2005). Many proteins with either drug or vaccine application require glycosylation, which cannot be practically produced in prokaryotes or lower eukaryotes (Lisowska, 2002). Therefore, the capability of mammalian cells to correctly glycosylate target protein makes this approach one of the most attractive methods for the production of recombinant biologically active products. However, whilst the use of mammalian cells such as Chinese hamster ovaries (CHO) or human embryonic kidney (HEK)-293 is well documented, the process of creating stable mammalian cell lines can often be laborious and time consuming (Hunt, 2005). Moreover, expression with mammalian cell systems typically produces substantially lower yields compared with other expression systems. Owing to the low yield, greater cost is required to produce adequate amounts of protein for subsequent application. Transient expression systems that permits milligram to gram quantities of protein, by utilising suspension cell lines, provide a viable alternative (Wurm and Bernard, 1999). Transfection of mammalian cells (with plasmid- or virus-based vectors) is often used for rapidly expressing high amounts of protein for laboratory scale application or preliminary evaluation of drug or vaccine targets and can be further enhanced with the use of bio-reactors.

4.1.1.5 Plants as expression system

“Molecular farming” in plants offers considerable advantages in cost and safety and has therefore received increasing attention as a promising platform for the production of complex recombinant vaccines and drugs (Basaran and Rodriguez-Cerezo, 2008; Lau and Sun, 2009; Tremblay *et al.*, 2010). Plants, as eukaryotic organisms, hold all the features to successfully express, fold and post-translationally modify foreign protein. Over the last 3 decades a range of different plant and vector systems have been utilised

for the production of a list of antigens that includes viral (Mason *et al.*, 1996; Kalthoff *et al.*, 2010), bacterial (Brennan *et al.*, 1999; Rigano *et al.*, 2003) and auto-immune antigens (Fitchen *et al.*, 1995).

Several inducible expression systems have been developed for temporal and spatial transgene expression in plants and have to date mainly been employed in tobacco, rice, Arabidopsis, tomato and maize (Borghi, 2010). Though most recombinant vaccine and therapeutic antibodies currently in preclinical and clinical trials were produced in whole-plant expression systems, plant cell suspension cultures (which integrates the characteristics of whole-plant systems with those of microbial systems) hold great potential as alternative "factories" (Xu *et al.*, 2011). The most frequently used host lines include the fast growing, robust and easily transformed tobacco *Bright yellow* (BY)-2 and *Nicotiana tabacum* (NT)-1 lines. However, low protein productivity (0.01 – 200 mg/L) within these systems remains a major obstacle, slowing the commercialisation of plant cells as bio-production platforms (Xu *et al.*, 2011). Approaches to overcome and solve the associated challenges, including non-mammalian glycosylation and genetic instability, involve several molecular approaches. To enhance gene transcription, strong constitutive or inducible promoters can be implemented in the transgene vector systems (Desai *et al.*, 2010). Translational efficiency can be improved by optimising the recombinant gene codon usage (Gustafsson *et al.*, 2004). To minimise post-translation degradation by proteases, foreign protein can be targeted to sub-cellular compartments such as the ER where proteolysis is less likely to occur (Sharma and Sharma, 2009; Desai *et al.*, 2010). Finally, to stabilise and allow enhanced secretion of a target protein, the protein can be expressed together with a fusion partner such as a highly-expressed and stable plant protein or a fusion tag (e.g. Hyp-Glycomodule) (Streatfield, 2007; Xu *et al.*, 2011).

Of great interest, is the recent concept of edible vaccines, derived from plants in which the antigen was recombinantly expressed. Transgenic plants are an ideal means by which to produce oral vaccines, not only because of the molecular characteristics of plants but also the due to their anatomic features such as the rigid cell walls that enables protection of the antigenic protein from the acidic environment of the stomach (Rigano and Walmsley, 2005). The feasibility of this approach have been demonstrated for Hepatitis-B in human and mice trials (Yusibov *et al.*, 2011) and have already been extended to malaria (Birkholtz *et al.*, 2008) and tick control (Bensaci *et al.*, 2012).

However, before edible plant vaccines can serve as a viable control strategy, it is obvious that studies investigating the formulation, delivery, antigen stability and immunogenicity are required.

4.1.1.6 Cell-free expression systems

With *in vivo* and *in vitro* gene expression technologies having problems associated with living cells such as protein degradation and aggregation, loss of template and laborious experimental steps, cell-free expression represents a promising alternative. Within a cell-free expression system an optimal environment for protein synthesis is created, since all metabolic resources are assigned to the particular protein to be synthesised (Vinarov *et al.*, 2006). Cell-free expression allow for the expression of difficult proteins (such as complex conformational proteins, cytotoxic proteins and proteins that usually undergo proteolysis) in a controlled environment without intact cells (Shimizu *et al.*, 2006).

Cell-free systems are based on the cellular ribosomal protein synthesis system. Thus, for efficient protein synthesis within a cell-free system several factors are compulsory. These include: transcription of mRNA, aminoacyl tRNA, energy provision and translation factors (Shimizu *et al.*, 2006). The most widely used open expression systems are bacterial S30 lysate fraction, wheat germ extract and rabbit reticulocyte lysates. The cell extracts from these cells containing the machinery necessary for protein production (including ribosomes, translational factors, aminoacyl tRNA synthase and tRNAs) serve as basis for protein synthesis. Once the gene of interest, cloned into a DNA plasmid vector, is transcribed by RNA polymerases (e.g. T7 and Sp6) the cell extracts containing the necessary requirements for translation is added and transcribed product is translated into protein. However, in most cases for effective recombinant protein production supplementation of the S30 lysate with additional energy regenerating enzymes and their substrates (e.g. phosphoenolpyruvate (PEP)/ phosphoenolpyruvate kinase (PK) and creatine phosphate (CP)/ creatine kinase (CK)) are required (Shimizu *et al.*, 2006). For cell-free systems to express complex proteins in their correctly folded and functional active state, additional components such as chaperones may also be required (Bukau *et al.*, 2000; Hartl and Hayer-Hartl, 2002). Since no cell disruption is required purification is simple. However, to maintain the

native conformation state during purification proteins can first be transferred by means of a gradient sift into liposomes (Kaneda *et al.*, 2009).

Since Spirin *et al.* demonstrated in 1988 that cell-free translation could be utilised as a tool for producing proteins, several optimisation strategies (e.g. improved cell lysate preparation and development of energy regeneration systems) have been implemented, increasing productivity up to several milligram quantities (Spirin *et al.*, 1988; Michel and Wüthrich, 2012). This technology therefore has become a viable platform for high-throughput protein expression for both laboratory and pharmaceutical use (Casteleijn *et al.*, 2013). Its controlled environment also provides a means to study complex biological processes and holds the potential for protein engineering, allowing the development of proteins with novel functions and the synthesis of unnatural amino acids (Shimizu *et al.*, 2006). However, broad application of this expression method is mainly altered by the great costs involved to mimic the correct protein synthesis environment.

4.1.2. Tick protein expression

Purification of sufficient amounts of candidate tick vaccine antigens has been one of the major limitations to the implementation of host vaccination as an alternative tick control method. With the onset of vaccines as control strategy against ticks, host animals were initially vaccinated with extracts of the parasite which yielded only moderate success (Elvin and Kemp, 1994). Isolation and purification of tick antigens, with the ultimate aim of preparing efficacious antigens for large scale vaccination trails, has proved to be viable but remains difficult and time-consuming (Willadsen *et al.*, 1988). Together with the immense amount of sequencing data generated, scientific and technological advances offered in the post genomic era, has resulted in a paradigm shift of vaccine development (Oberg *et al.*, 2011). Reverse vaccinology and reverse genetics approaches now allow for the identification and characterisation of anti-tick targets, and recombinant expression permit the production of large amounts of such proteins.

To date most identified and evaluated tick antigens have been expressed utilising *E. coli* as expression system (Table 4.1). These include the recent work of Canales *et al*. (2008a), where they demonstrated that a recombinant protein comprising the BM95 immunogenic peptides fused to the *A. marginale* MSP1α N-terminal region (BM95-

MSP1α) can be expressed on the *E. coli* membrane, providing a simple and cost-effective process for the production of vaccine preparation.

Table 4.1. Summary of a selection of tick proteins that have been expressed.

Species	Antigen name	Expression host	Reference
R. microplus	Bm86 & Bm91	*E. coli*	(Rand *et al.*, 1989; Willadsen *et al.*, 1996)
	Cysteine protease	*E. coli*	(Renard *et al.*, 2000)
	Bm95	*E. coli* (MSP1α)	(Canales *et al.*, 2008a)
	5'-Nucleotidase	Baculovirus	(Liyou *et al.*, 1999)
	Bm86	Baculovirus	(Richardson *et al.*, 1993)
	Bm86	Fungel cells (*Aspergillus nidulans*)	(Turnbull *et al.*, 1990)
	Bm86 &Bm95	*P. pastoris*	(Rodriguez *et al.*, 1994)
R. annulatus	Glutathione S-transferase	*E. coli*	(da Silva Vaz *et al.*, 2004)
Hyalomma anatolicum anatolicum	Haa86	*P. pastoris*	(Azhahianambi *et al.*, 2009)
R. appendiculatus	64 TRP1	*E. coli*	(Trimnell *et al.*, 2002)
	64 TRP2	*E. coli*	
	64 TRP3	*E. coli*	
	64 TRP4	*E. coli*	
	64 TRP5	*E. coli*	
	64 TRP6	*E. coli*	
	RAS-3, RAS-4 & RIM36	*E. coli*	(Imamura *et al.*, 2009)
H. longicornis	RH50	*E. coli*	(Zhou *et al.*, 2006a)
	P29	*E. coli*	(Mulenga *et al.*, 1999)
	HL34	*E. coli*	(Tsuda *et al.*, 2001)
	HLS2	*E. coli*	(Imamura *et al.*, 2005)
	HLS1	*E. coli*	(Sugino *et al.*, 2003)
	Leucine amino peptidase	*E. coli*	(Hatta *et al.*, 2006)
	Serine protease	*E. coli*	(Anisuzzaman *et al.*, 2012)
	Chitinase	Baculovirus	(Assenga *et al.*, 2006)
Haemaphysalis qinghaiensis	HqCRT	*E. coli*	(Gao *et al.*, 2008a)
	HqTnT	*E. coli*	(Gao *et al.*, 2008b)
I. ricinus	Metis 1	Mammalian cells	(Decrem *et al.*, 2007)
	Metis 2	Mammalian cells	
Ornithodoros savignyi	Apyrase	*P. pastoris*	(Stutzer *et al.*, 2009)

With Bm86 identified as the highest efficacious tick antigen, a large amount of studies have evidently focused on expressing Bm86 (and orthologous), in a conformation as close possible to the native protein and in as high quantities as possible for subsequent

vaccination. Of the first vaccination studies conducted with recombinant Bm86 was performed utilising β-galactosidase fusion proteins containing most or all of the Bm86 amino acids which were expressed as insoluble inclusion bodies by *E. coli* (Rand *et al.*, 1989). A near full length Bm86 has also been expressed in fungal cells (*Aspergillus nidulans*) and in *Spodoptera frugiperda* (Sf9) insect cells using the baculovirus system (Turnbull *et al.*, 1990; Richardson *et al.*, 1993). The expression yields in the baculovirus system were very high, the protein was successfully secreted and glycosylated and subsequently showed strong protection against *R. microplus* in vaccination trails (Elvin and Kemp, 1994). For complex proteins, such as the membrane bound gut glycoprotein Bm86, it has been proven that more advance expression systems (e.g. baculovirus-insect cell systems and *P. pastoris*) where post-translational modifications (e.g. glycosylation, disulphide bond formation and protein phosphorylation) takes place, are more efficient in expressing secreted full length protein (Rodriguez *et al.*, 1994; Canales *et al.*, 2008b). It has furthermore been shown that *P. pastoris* are likely to produce recombinant protein in a conformation closely resembling the native protein, which in turn may increase the immunogenic and protective properties of an antigen (García-García *et al.*, 1998). Therefore, several studies of tick antigen recombinant protein production (including the largescale production of GAVAC^{TM}) have focused on using *P. pastoris* for optimal expression, to finally increase vaccination efficiency. However, *P. pastoris* expression in comparison to a simple bacterial expression might be more difficult and require time consuming optimisation, with the yeast system having limitations such as inconsistent glycosylation, excessive protease activity over long expression periods and the limited selectable markers available for transformation (Balamurugan *et al.*, 2007).

Difficulties with production of full length functional tick proteins are a serious obstacle, and therefore vaccination with truncated tick proteins also commenced. The proof of principle that more than one region of a tick protein molecule is protective was demonstrated by Tellam *et al.* (1992), which showed that truncated constructs of the Bm86 used in vaccine trails raised an effective protective immune response. In the broad vaccine development field (including vaccines against parasitic and infectious diseases) considerable success has been made with peptide based vaccines. However, finding an optimal platform with all the necessary requirements for an effective and specific immune response remains a major challenge. The ongoing development of more bio-informatic tools and reverse genetics approaches in tick research, coupled

with more genome sequence data becoming available, provide the primary tools for identifying and characterising vaccine targets and immunogenic peptides. Ultimately, for a comprehensive understanding of a particular target, several heterologous expression systems permit the production of recombinant protein. Over time these systems will continue to develop and new ones will arise, all to aim for the most successful production of tick proteins, especially for use as recombinant vaccines. Alternative expression systems, which hold great potential, include the use of plants as expression systems (i.e. edible vaccines) as well as the use of chimeric constructs in simple expression host cells.

4.1.3. Chimeric fusion enhances recombinant expression and vaccine immunogenicity

Gene fusion (i.e. production of chimeric proteins) is a viable strategy to circumvent many of the problems associated with recombinant protein production, especially when *E. coli* is used as expression host. Additionally, fusion of different immunogenic peptides has been shown to enhance immunogenicity (Patarroyo *et al.*, 2002; Tan *et al.*, 2011). Therefore, gene fusion technologies continued over the years to expand, with the introduction of new fusion partners, purification and detection tags, cleavage reagents and ways to display peptides on the surface of the expression host cell.

For example, to enhance solubility of proteins expressed in *E. coli* an adaptable expression system based on the *E. coli* thioredoxin (trxA) as gene fusion partner, has been developed (LaVallie *et al.*, 1993): Proteins which are initially expressed in *E. coli* as insoluble or in inclusion bodies, are soluble as thioredoxin fusions and biologically active. On the other hand, expression of recombinant membrane proteins or exoplasmic domains presents another set of distinct challenges. Therefore these proteins are often expressed using mammalian expression systems, which are expensive and yield low productivity, which in turn is not optimal for cost effective large scale vaccine production. To allow for inexpensive, high yield expression of exoplasmic domains of tick membrane proteins Canales *et al.* (2008a) established an *E. coli* membrane expression system which utilises the N-terminal of a major surface antigen present on *Anaplasma marginale* (MSP-1) (see section 4.1.4 for full description).

The approach of utilising vaccines encoding multiple epitope-peptides to improve immunogenicity has been widely investigated and implemented. A recent example is the finding that an enhanced protection response can be obtained against influenza by using a chimeric vaccine (Tan *et al.*, 2011). The vaccine is based on the matrix protein 2, conserved throughout inlfluenza A viruses, fused to norovirus P particle. The vaccine therefore not only holds enhanced protection against influenza but also permits protection against noroviruses.

Malaria parasite studies have reported the design and expression of *Plasmodium yoelli* chimeric recombinant proteins as an effective platform for malaria vaccines (Shi *et al.*, 2007). Shi *et al.* (2007) demonstrated that immunisation with a designed chimeric antigen construct, of the epitope rich regions of the merozoite surface protein-1 and 8, increased overall protection. Furthermore, in late studies the potential of a synergistic or additive effect by creating a multistage vaccine consisting of more than one chimeric antigen construct, has been evaluated (Singh *et al.*, 2012). The multistage vaccine was successfully expressed in *E. coli* and induced a robust immune response against each of the individual components.

The use of chimeric antigen constructs in tick vaccine development has also been explored and demonstrated. Patarroyo *et al.* (2002) synthetically constructed a continuous chimeric construct of 3 peptides (SBm4912, SBm7462 and SBm19733) from different regions of Bm86. With cattle vaccination it was evident that a more efficient immune response against the cattle tick *R. microplus* was observed, compared to vaccination with full length Bm86 (Patarroyo *et al.*, 2002). This set the grounds that chimeric antigen constructs holds great potential in tick vaccine development.

4.1.4 A novel MSP1α-based on-membrane expression system

The novel membrane expression system, developed by Canales *et al.* (2008a), permits expression of proteins or/and peptides on the *E. coli* membrane by expressing the gene of interest in a specialised vector that allows the fusion to the N-terminal domain of mutated MSP1α (Figure 4.6). This novel expression was first constructed by cloning a mutated MSP1α gene into the bacterial FLAG® expression vector (de la Fuente *et al.*, 2001). The FLAG® *E. coli* expression vector confer ampicillin resistance for easy

selection and the construct is under control of the strong tac promoter (a hybrid of the trp and lac promoters). This allow the production of protein expression levels in excess of 10 mg/L. The mutated MSP1α is derived from the major surface protein of *A. marginale*. By means of a variable number of tandem 28 or 29 amino acid repeats located in the N-terminal of the protein that contains an adhesion domain permitting successful infection of cells, native MSP1α is involved in the adhesion of the pathogen to tick cells and bovine erythrocytes (de la Fuente *et al.*, 2004). The mutant MSP1α was constructed by removing the tandem repeats, leaving an exposed N-terminal domain which is surface displayed when the protein is expressed in *E. coli*. Once the mutated MSP1α was cloned into pFLAG expression vector (pAFOR1) any gene of interest could be cloned upstream of MSP1α and upon expression will be transcribed, translated and surface displayed as a recombinant chimeric protein.

The proof of principle was demonstrated by Canales *et al.* (2008a) who expressed immunogenic peptides of Bm95 fused to the N-terminal of *A. mariginale* MSP1α for presentation on the *E. coli* membrane. Expressed fraction analysis proved that chimeric protein was expressed and surface exposed as well as recoginised by anti-Bm86 and anti-MSP1α antibodies. Subsequent vaccine trails were done in rabbits (Canales *et al.*, 2009) and cattle (Almazán *et al.*, 2012) validating that bacterial membrane extracts containing chimeric protein are protective against tick infestation.

In subsequent studies the use of this system was extended. Almazán *et al*. (2012) not only expressed Bm95, but also subolesin (SUB), elongation factor 1α (EF1α) and ubiquitin using this approach. In cattle trials a vaccine efficacy of more than 60% was observed for the fusion protein SUB-MSP1α against *R. microplus* and *R. annulatus*. The Bm95-MSP1α vaccine resulted in the highest reduction of tick numbers while SUB-MSP1α showed most effect on egg fertility compared to the other tested vaccine candidates (Almazán *et al.*, 2012).

This novel approach provides a simple, cost effective and rapid means for the production of tick protective antigens by using an antigenic protein chimera which is surface displayed on *E. coli* cells. Since previous vaccination studies have shown that MSP1α in itself has good protection efficiency against anaplamosis (de la Fuente *et al.*, 2002) a recombinant vaccine based on this approach therefore not only holds the promise of enhanced protection against the tick vector but also permits protection

against anaplasmosis. A further key advantage of this system is that the recombinant bacterial membrane fractions can be directly used for vaccination, as it has been observed that bacterial components have an added adjuvant effect in vaccine preparations (Canales *et al.*, 2008a). Even though *E. coli* is a known pathogen in animal hosts, no pathogen related illness or any other negative health effects was observed in the studies conducted so far on rabbit and cattle hosts (Canales *et al.*, 2008a; Canales *et al.*, 2009; Almazán *et al.*, 2012). Finally, by fusing tick antigens to MSP1α high expression levels (between 3 – 12 % of the total cell protein) were obtained (Almazán *et al.*, 2012). Altogether this expression method holds promise potential for the production of recombinant anti-tick vaccine (de la Fuente *et al.,* 2007; Guerrero, *et al.,* 2006). The use of bacterial membrane fractions will also greatly reduce vaccine production cost for proof-of-principle first round trails, since no costly protein purification is needed. This method is fundamental for first phase analysis of an antigen, which will be followed by expression and vaccine optimisation and subsequent second round trials, if the antigen showed a significant immune response.

4.1.5 Hypothesis

The MSP1 α- *E. coli* expression system can be utilised to express selected domains of the three promising metzincin antigens, suitable for first cattle vaccination trials.

4.1.6 Aims

- Amplification and cloning of full length BmMP1, BmMP2 and As51 into pGEM®-T easy vector.
- To reconstruct the *E. coli* expression vector pMBAXF3 (MSP1α-Bm95) to pAFOR1x vector, by removing Bm95 as well as reconstructing and optimising the *Xho*I and *EcoR*I restriction enzyme sites.
- To determine the most immunogentic (epitope rich) ~ 150 amino acid domain of BmMP1, BmMP2 and As51, using *in silico* tools.
- Directional cloning of selected domains of the three *R. microplus* metzinins into the MSP1α expression vector pAFOR1x.

- Expression of chimeric MSP1α-metzincin recombinant proteins under optimised temperatures and time frames.
- Amino acid-sequence validation of the recombinantly expressed proteins, using LC-MS-MS.

4.2 Materials and Methods

4.2.1 Cloning of full length BmMP1, BmMP2 and As51 into pGEM®-T Easy vector

4.2.1.1 PCR amplification and purification

In order to amplify and clone each of the three *R. microplus* metzincin transcripts, gene specific primers (GSPs) amplifying the entire known coding sequence of BmMP1, BmMP2 and As51 were designed. Single-stranded cDNA synthesised from *R. microplus* mixed lifestages was used as template (section 2.2.5). For each reaction 1 µl of cDNA template was combined with 10 pmol forward and 10 pmol reverse gene specific primer, 1x TaKaRa ExTaq Reaction Buffer™ (25 mM TAPS pH 9.3 at 25°C; 50 mM KCl; 2 mM MgCl2; 1 mM 2-mercaptoethanol), 0.2 mM of each dNTP, 1.25 U TaKaRa ExTaq™ (Takara, Japan) and water to a final volume of 20 µl. Reactions were carried out in a 2720 Thermal Cycler (Applied Biosystems, USA). Briefly, cDNA was denatured at 94°C for 2 minutes, before enzyme was added. Amplification consisted of 30 cycles of denaturation (94°C, 30 sec), annealing (at the Tm of the gene specific primers, 30 sec) and extension (72°C, 2 minutes), followed by a final extension step (72°C, 7 minutes). PCR products were analysed on a 2% w/v agarose/TAE gel as described in section 2.2.8, except instead of 50 ng EtBr.

To purify the PCR products from unincorporated nucleotides, remaining polymerase, primers and salts, PCR reactions were subjected to silica-column purification using the QIAquick PCR Purification Kit (Qiagen, USA). In brief, 5 volumes of Buffer PB (composition not supplied by manufacturers) were added to 1 volume of the PCR sample and were mixed. The sample was loaded on to a column and centrifuged for 30 sec at 13 500 x g. Once the flow through was discarded, the column was washed with 750 µl Buffer PE (composition not supplied by manufacturers) and centrifuged for 30 sec at 13 500 x g. Residual ethanol from Buffer PE was completely removed by centrifugating for an additional minute at 13 500 x g. To elute the DNA, the column was placed in a clean 1.5 ml tube, 30 µl elution buffer (10 mM Tris-HCl, pH 8.5) was loaded onto the center of the column and the column was incubated in a 37°C shaking

incubator for 5 minutes and subsequently centrifuged at 13 500 x g for 30 sec. The concentrations of the purified products were determined spectrophotometrically.

4.2.1.2 Ligation and selection of recombinant constructs (for a full description of each step refer to section 2.2.11)

All of the full length amplified products were ligated into the pGEM® - T Easy vector system (Promega, USA), as described in section 2.2.11.1. Briefly, ligation was performed at 16°C overnight with 1 µl T4 DNA Ligase (3 U/µl), 1 µl 10x Ligase buffer (Promega, USA) and an additional 1 µl ATP (0.1 pmol). Once the ligation mixture of each transcript was heat inactivated (70°C for 10 min), precipitated and resuspended, electrocompetent JM 109 *E. coli* cells (section 2.2.11.2) were transformed with recombinant plasmid via electroporation (section 2.2.11.3). For each transformant three LB-agar plates (2% w/v agar in LB-Broth), containing 50 µg/ml ampicillin, were prepared onto which 50 µl of a 1/10 dilution, undiluted or a resuspended pellet of transformed cells were plated. Plates were incubated upside down at 37°C over night and recombinant clones were identified with PCR colony screening, using the Sp6 and T7 primers (Table 2.3, section 2.2.11.4).

Recombinant plasmid for sequencing analysis and subsequent directional cloning was isolated using the Zyppy™ Miniprep Kit (section 2.2.11.5). DNA nucleotide sequences of recombinant clones were determined by automated nucleotide sequencing using the ABI PRISM® DT3100 Genetic Analyzer (PE Applied Biosystems, USA, supplied by the DNA Sequencing Facility of the Faculty of Natural and Agricultural Sciences at University of Pretoria (http://seqserve.bi.up.ac.za)) and the Big Dye version 3.1 sequencing kit (Perkin Elmer, USA). DNA sequences were analysed using the Bio Edit version 5.0.9 program (Hall, 1999). Once the results were confirmed and correlated by visual inspection of the electropherograms the nucleotide and amino acid sequences were aligned against their original astacin or reprolysin nucleotide sequences using Clustal W (http://www.ebi.ac.uk/clustlew/) (Thompson *et al.*, 1997).

4.2.2 Construction of the pAFOR1x expression vector

The plasmid pAF0R1 was initially constructed to express a mutant *A. marginale* MSP1α that lacks the tandem 23-31 amino acid repeats in the N-terminal surface-exposed region (de la Fuente *et al.*, 2003). We were fortunate to receive JM 109 cell stock containing the pAFOR1- Bm95 (pMBAX) chimera construct from our collaborators (Canales *et al.*, 2008a). To establish the exact construct that we received the first step was to sequence the construct. Briefly, 100 µl of JM109 cell stock (containing the pAFOR1- Bm95 construct) were inoculated and re-grown overnight at 37°C with shaking in 5 ml LB-ampicillin in 50 ml flasks. The plasmid was isolated utilising the Zyppy™ Miniprep Kit (Zymo research Corporation, USA) and submitted for sequencing.

In order to convert pMBAX back to pAFOR1 to express selected domains of BmMP1, BmMP2 and As51 as chimeras with the modified MSP1α, pMBAX was reconstructed to pAFOR1x by removing BM95 and inserting additional basepairs to regain the function of the *Xho*I and *EcoR*I restriction enzyme sites (Figure 4.6).

4.2.2.1 Excision of Bm95 from pMBAX

In order to remove Bm95 from the construct, the pMBAX plasmid was subjected to restriction enzyme digestion with *Xho*I and *Eco*RI. Digestion was performed in a reaction mixture containing 1 µg of plasmid, 0.5 µl *Xho*I (10 U/µl, Promega, USA), 0.5 µl *Eco*RI (12 U/µl, Promega, USA), 2 µl 10x reaction buffer D (6 mM Tris-HCl, 6 mM $MgCl_2$, 150 mM NaCl, pH 7.9), 0.2 µl acetylated BSA (10 µg/ul) and double distilled de-ionised water to a final volume of 20 µl. The reaction was incubated at 37°C for 4 hours, after which the enzyme was inactivated with incubation at 65°C for 15 minutes.

To establish if the BM95 fragment was successfully removed and to obtain the linearised product (pFLAG-CTC-MPS1a/ pAFOR1) DNA gel electrophoresis was conducted using the FlashGel^{TM}system. This system utilises a proprietary stain that is 5-20 times more sensitive than EtBr stain, allowing the detection of DNA levels as low as 0.10 ng per band. The main advantage of the system is that separation of DNA fragments can be monitored in real time and fragments can be recovered on the bench

without the use of UV illumination or the need of band excisions. In brief, 6 x 10 µl (500 ng) of the digested product was loaded with 2 µl 1x FlashGel™ Loading Dye in 6 respective wells on a 2.2% FlashGel™ Recovary Cassette alongside FlashGel™ DNA marker. Given that the system was designed for fast, high voltage separation the gel was ran at 275 V, separating the 146 bp (Bm95) band from the 6.9 kbp (pFLAG-CTC-MSP1α/ pAFOR1) band within a few minutes. Once it was clear that the Bm95 fragment was removed, the run was continued while monitoring migration of the linearised 6.9 kbp samples to be recovered. Just prior to the samples reaching the recovery wells, the power supply was turned off and disconnected. Excess running buffer was blotted from the recovery wells and 20 µl of FlashGel™ Recovery Buffer was added. The power supply was then reconnected and restarted. Once the bands to be recovered had entered the recovery wells, the power was turned off, and the recovery buffer of all 6 wells, containing the DNA, was removed using a pipette and pooled together. Once removed in the FlashGel™ Recovery Buffer the pooled sample was compatible with all downstream applications and the concentration was determined spectrophotometrically.

4.2.2.2 Amplification for the incorporation of *Xho*I and *Eco*RI cleavage sites

A set of gene specific primers (GSPs) was designed to regain and optimise the *Xho*I and *Eco*RI restriction enzyme sites' functionality to finally construct the expression vector pAFOR1x. The forward primer contains the partial *Eco*RI cleavage site followed by the MSP1α N-terminal gene specific area. The reverse primer contains the final base to complete the *Eco*RI cleavage site, the *Xho*I cleavage site and additional bases which flanks both cleavage sites for optimal cleavage efficiencies (Table 4.2).

Table 4.2. Characteristics of the primers used for PCR amplification and incorporation of restriction enzyme sites of the pAFOR1x vector. The nucleotide bases to construct the *Eco*RI site are indicated in yellow and the *Xho*I site in magenta. The start codon of MSP1α is indicated in green, followed by the N-terminal gene specific area of MSP1α in turquoise. Additional bases to maintain the construct in frame and assure optimal restriction enzyme cleavage are indicated in red and the pFLAG-CTC vector in grey.

Primer Name	Primer sequence (5' → 3')	Tm (°C)
LongAFOR1Fw	AAT TCC ATG TTA GCG GCT AAT TGG CG	64.5
LongAFOR1Rvnew	CCC CTC GAG AAG CTT CAT ATG ATA TCT CC	67.45

For amplification *Pfu* DNA polymerase was used to generate blunt ends upon amplification. Fifty nanograms of template was added to a reaction solution containing 10 pmol forward and 10 pmol reverse gene specific primer and final concentrations of 0.2 mM of each dNTP, 1x *Pfu* Buffer and 1 U of *Pfu* DNA polymerase (Promega, USA). Amplification was achieved after an initial denaturing step (95 °C, 3 minutes), followed by 35 cycles of denaturation (95°C, 30 sec), annealing (63°C , 30 sec) and extension (68°C, 5 min), followed by a final extension step (68°C, 7 minutes). PCR product was analysed on a 1% w/v agarose/TAE/GelRed gel. Largescale PCR reactions were performed under the same conditions for subsequent amplification of sufficient amounts of product to serve as vector during expression. Modified amplicon was purified from the gel using the Promega Wizard® SV Gel and PCR clean-up system (Promega, USA) and concentrations were determined spectrophotometrically.

Briefly, the DNA fragment of interest was excised from the gel in a minimal volume of agarose. The gel slice was transferred to a weighed 1.5 ml microcentrifuge tube and the weight was recorded. Membrane binding solution (4.5 M guanidine isothiocyanate, 0.5 M potassium acetate, pH 5.0) was added at a ratio of 10 µl of solution per 10 mg of agarose gel slice. The mixture was vortexed and incubated at 50–65°C until the gel slice was completely dissolved, transferred to a SV Minicolumn and incubated for a minute at room temperature before the column was centrifuged at 16,000 × *g* for 1 minute. The flow-through was discarded and the column was washed by adding 700 µl of Membrane wash solution (10 mM potassium acetate (pH 5.0), 80% ethanol, 16.7µM EDTA, pH 8.0) and subsequent centrifugation at 16,000 × *g* for 1 minute. The wash step was repeated with 500 µl of Membrane Wash Solution and centrifugation for 5 minutes at 16,000 × *g*, with an additional centrifugation for 1 minute after the flow-through was removed to allow evaporation of any residual ethanol. Finally, the column was transferred to a clean 1.5 ml microcentrifuge tube, 50 µl of nuclease-free water was applied directly to the center of the column, it was incubated at room temperature for 1 minute and centrifuge for 1 minute at 16,000 × *g* to obtain the eluted DNA. The concentration was determined spectrophotometrically and the purified linear pAFOR1x product was stored at –20°C.

4.2.2.3 Blunt-end ligation and transformation

For blunt-end ligation of pAFOR1x the sample was first subjected to phosphorylation by T4 polynucleotide kinase (T4 PNK). This enzyme catalyses the transfer of the γ-phosphate from ATP to the 5´-terminus of polynucleotides (Richardson, 1965). Briefly, 10 µl of sample (containing 0.2 pmol of 5'-OH ends) was added to a reaction containing 1 x kinase buffer, 10 U T4 PNK (Promega, USA), 0.2 pmol ATP and double distilled deionised water to a final volume of 20 µl. The reaction was mixed briefly by centrifugation, incubated at 37°C for 30 min and finally the reaction was haltered by snap cooling on ice water. Once phosphorylation was completed, ligation was performed at 16°C overnight with 3 U T4 DNA Ligase and 1x Ligase buffer (Promega, USA). The reaction was heat inactivated by incubation at 70°C for 15 minutes and precipitated by means of standard NaAcetate/ethanol precipitation (section 2.2.12). The ligation reaction was transformed into electrocompetent JM 109 *E. coli* cells by means of electroporation (section 2.2.11.3). Three LB-agar plates (2% w/v agar in LB-Broth), containing 50 µg/ml ampicillin, were prepared onto which 50 µl of a 1/10 dilution, undiluted or a resuspended pellet of transformed cells were plated. Plates were incubated upside down at 37°C over night. Selected colonies were re-grown overnight at 37°C with shaking in 5 ml LB-ampicillin in 50 ml flasks. To asses if the linear phosphorylated product successfully self-ligated to form pAFOR1x, plasmid for restriction enzyme digestion analysis was isolated using the Zyppy™ Miniprep Kit (Zymo research Corporation, USA).

4.2.2.4 Restriction enzyme digestion and sequence analysis

Isolated plasmid (pAFOR1x) was submitted to restriction enzyme digestion to establish if the reconstruction of the *Xho*I and *Eco*RI cleavage sites was successful. Two separate digestions were performed, with each reaction containing 1 µg of plasmid, 1x NEB buffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol, pH 7.9), 0.2 µl acetylated BSA (10 µg/ul) and 20 U of *Xho*I (New England Biolabs, Inc., USA) and *Eco*RI (New England Biolabs, Inc., USA), respectively. The reaction volume was adjusted with double distilled de-ionised water to a final volume of 20 µl and incubated at 37°C for 4 hours, after which the enzyme was

inactivated with incubation at 65°C for 15 minutes. Products were analysed on a 1% w/v agarose/TAE/EtBr gel along non-digested plasmid. Finally, DNA nucleotide sequences of the newly constructed expression vector, pAFOR1x was determined with automated nucleotide sequencing conducted by Inqaba Biotec (Pretoria, RSA).

4.2.3 Domain selection

Expression in the MSP1α – expression system allows insertion of transcripts only up to 450 basepairs (150 amino acids). Therefore, for each of the 3 metzincins the most epitope rich 150 amino acid domain was selected using a variety of *in silico* tools. Criteria included prediction of B- and T-cell epitopes, surface accessibility and secondary structures. To predict antigen binding to MHC class I and II, the top 10 binders to 3 HLA-DRB1 alleles for each metzincin were identified using MHCPred (http://jenner.ac.uk/MHCPred) (Guan *et al.*, 2003). Linear B-cell epitopes were predicted using BepiPred (http://www.cbs.dtu.dk/services/BepiPred) (Levitt, 1978; Larsen *et al.*, 2006). ProPred (http://www.imtech.res.in/raghava/propred) was used to predict MHC Class-II binding regions in the antigen sequences, using quantitative matrices derived from published literature (Sturniolo *et al.*, 1999; Singh and Raghava, 2002). ProPred assists in locating promiscuous binding regions within an antigen that are useful in selecting vaccine candidates. A propensity scale method that incorporates hydrophilicity, surface accessibility and flexibility of each amino acid in order to predict immunogenicity, designed by Kolaskar and Tongaonker, was used (Kolaskar and Tongaonkar, 1990). Furthermore, to assess the surface accessibility across each protein, the Emini method that incorporates the frequency of amino acid residues in surface-accessible areas was used (Janin and Wodak, 1978; Emini *et al.*, 1985). The presence of β-turns was detected using a program that makes use of artificial neural networks (Kaur and Raghava, 2004). For each program, each amino acid that was part of a good predicted binder or β-turn received a point. Protein sequences were evaluated for areas that were recognised by the majority of the predictors, in order to select a 150 amino acid domain for each metzincin for expression.

Secondly, the antigenicity of the selected domains were ranked and compared to their respective VaxiJen (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) scores. This program predicts potential antigens in an alignment independent way, by assigning

a score to each selected domain, with the highest scores predicting best antigen. The threshold cut-off value for being a probable antigen based on the parasite model of VaxiJen is set at 0.5.

4.2.4 Directional sub-cloning into pAFOR1x

4.2.4.1 Primer design and insert preparation for sub-cloning of the selected domains

For expression, the selected 150 amino acid metzincin domains needed to be cloned inframe, upstream of the exoplasmic N-terminal domain of MSP1α (downstream of the pFLAG-CTC starting codon). A set of gene specific primers were designed for the sub-cloning of the selected domains. Each of the forward primers contain an *Eco*RI cleavage site. Reverse primers contain an *Xho*I restriction enzyme site and an entrokinase site to allow subsequent cleavage of the expressed domains from MSP1α, if needed.

For BmMP2 and As51 successfully constructed pGEM-plasmid served as template (section 4.2.1.2). For BmMP1 a sequence encoding the selected 150 amino acid domain was chemical synthesised and cloned into pUC57 by GenScript Inc. (USA). For each reaction 1 µl of a 1/1000 dilution of the respective purified plasmid served as template. This was combined with 10 pmol forward and 10 pmol reverse gene specific primer, 2x KAPATaq Ready Mix DNA Polymerase (KAPABiosystems, USA) and water to a final volume of 25 µl. Initial denaturation of the template was performed at 94°C for 2 minutes. Amplification consisted of 30 cycles of template denaturation (94°C, 30 sec), annealing (30 sec) and extension (72°C, 2 minutes), followed by a final extension step (72°C, 7 minutes). For the initial 2 cycles the annealing temperature was set at the Tm of the gene specific region of the primers, thereafter the annealing temperature was increased to that of the full length primer. The PCR products were analysed on a 2% w/v agarose/TAE/EtBr gel. Largescale PCR reactions were performed under the same conditions for subsequent amplification of sufficient amounts of product for restriction enzyme digestion and sub-cloning.

4.2.4.2 Preparation of the pAFOR1x plasmid (isolation and phosphorylation)

To obtain high concentrations of pure pAFOR1x plasmid for downstream application, the NucleoBond® Xtra Midi plasmid purification kit (Macherey-Nagel,Germany) was used. From overnight cultures (100 ml), two aliquots of 50 ml cells were collected by centrifugation at 4500 x g for 10 min (4 °C). The cell pellets were resuspended in 6 ml suspension buffer S1 (50 mM Tris-HCl, 10 mM EDTA, 100 µg/ml RNAse A, pH8) and subsequently 6 ml lysis buffer S2 (0.2 M NaOH, 1% SDS) was added. After the samples were incubated at room temperature for 5 minutes, 6 ml binding buffer S3 (2.8 M potassium acetate, pH 5.1) was added and the samples were incubated on ice for 5 minutes. Once cell debris were removed with the use of gravity filtration through filter paper, the filtrate were loaded onto a NucleoBond® filter column which have been pre-equilibrated in buffer N2 (100 mM Tris-H_3PO_4,15% ethanol, 900 mM KCl, 0.15% Triton X-100, pH 6.3). After the column was washed with 10 ml buffer N3 (100 mM Tris-H_3PO_4, 15% ethanol, 1.15 M KCl, pH 6.3) the DNA was eluted in 5 ml elution buffer N5 (100 mMTris-H_3PO_4 , 15% ethanol, 1 M KCl, pH 8.5). The sample was divided into five 1 ml fractions, 750 µl isopropanol was added and the reaction mixture was centrifuged for 45 minutes at 16000 x g (4°C). The supernatant was discarded and the precipitate washed with 500 µl 70% ethanol and subsequent centrifugation. The wash step was repeated twice, the pellet was dried *in vacuo* and dissolved in 50 µl double distilled deionised water. DNA concentrations were determined spectrophotometrically.

4.2.4.3 Directional cloning

The pAFOR1x plasmid and the 4 amplified metzincin domains were all subjected to restriction enzyme digestion with *Eco*RI and *Xho*I. For each reaction 1 µg DNA was combined with, 1x NEB buffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol, pH 7.9), 0.2 µl acetylated BSA (10 µg/ul), 20 U of *Xho*I (New England Biolabs, Inc., USA) and 20 U *Eco*RI (New England Biolabs, Inc., USA). The reaction volume was adjusted with double distilled de-ionised water to a final volume of 20 µl and incubated at 37°C (16 hours for plasmid and 4 hours for metzincin amplicons), after which the enzyme was inactivated with incubation at 65°C for 15 minutes. Digested products were purified using the Promega Wizard® SV Gel and PCR clean-up system (Promega, USA) (section 4.2.2.2). Products were eluted

in 30 µl elution buffer and the yields were determined spectrophotometrically. The digestion efficiency on the plasmid was evaluated by analysing 5 µl of the plasmid sample on a 1% w/v agarose/TAE/EtBr gel.

To circumvent self-ligation of the plasmid without insert, the digested pAFOR1x plasmid was subjected to dephosphorylation. Briefly, 5.5 µl shrimp alkaline phosphatase (1 U/µl) and 1 x reaction buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, pH 9.5) was added to 500 ng plasmid. The reaction was incubated at 37°C for 45 minutes and heat-inctivated at 65°C for 15 minutes.

To generate MSP1α-metzincin constructs the double digested amplicons were ligated at an insert:vector ratio of 3:1 utilising 50 ng of dephosphorylated pAFOR1x plasmid. Ligation was performed at 4°C (overnight) with 1 µl T4 DNA Ligase (3 U/µl), 1 x Ligase buffer (300 mM Tris-HCl, 100 mM $MgCl_2$, 100 mM DTT, 10 mM ATP, 10% PEG, pH 7.8). Once the ligation mixture of each transcript was heat inactivated, precipitated and resuspended, electrocompetent JM 109 *E. coli* cells were transformed via electroporation (section 2.2.11.3). For each transformant three LB-agar plates (2% w/v agar in LB-Broth), containing 50 µg/ml ampicillin, were prepared onto which 50 µl of a 1/10 dilution, undiluted or a resuspended pellet of transformed cells were plated. Plates were incubated upside down at 37°C over night and recombinant clones were identified with PCR colony screening (section 2.2.11.4), using forward and reverse vector specific primers (Table 4.3).

Clones displaying the correct insert sizes were selected and re-grown overnight at 37°C with shaking in 5 ml LB-ampicillin in 50 ml flasks. Cell stocks were made, from the overnight cultures in LB-Broth containing 10% v/v glycerol and stored at -70 °C. From the remaining culture, plasmid were isolated with the Zyppy™ Miniprep Kit (Zymo research Corporation, USA) and submitted for sequencing at Inqaba Biotec (Pretoria, RSA).

Table 4.3. Characteristics of the primers used for colony screening PCR to determine positive pAFOR1x-metzincin clones.

Primer Name	Primer sequence (5' → 3')	Tm (°C)
FwSeqMSP	CATCATAACGGTTCTGGCAAATATTC	61.4
MSP1SeqRev_new	GACTTCTGCAACACGCAGGGGAGC	69.6

4.2.5 Expression of MSP1α-metzincin chimeric protein in JM109 cells

Two positive clones for each of the 4 metzincin transcripts, native MSP1α (pAFOR1x alone, positive control) and native uninduced JM109 cells (negative control) were selected for expression. Cells (14 ml) from an overnight culture was inoculated and grown in 200 ml LB-broth (containing 50 µg/ml ampicilin and glucose to a final concentration of 0.4%) to an OD_{600nm} of 0.4-0.5. Expression was induced with isopropyl-d-thiogalactopyranoside (IPTG) (final concentration of 0.5mM). Additional glucose to a final concentration of 0.4% was furthermore added, to maintain a nutrient rich environment for the *E. coli*. Cultures were grown over 12 hours (200 rpm) at 37°C, 28°C and 20°C, respectively. Over a 24 hour period 15 ml aliquots were collected hourly.

Two milliliter cell aliquots of each time-point were harvested by centrifugation at 3800× g for 10 min (4 °C). Pellets were washed two times by adding 1.4 ml Tris-buffered saline (TBS, 20 mM Tris-HCl, 150 mM NaCl, pH 7.4) and 14 µl glucose (2 M) and subsequent centrifugation at 3800 × g for 10 min (4 °C). Once 1 ml of disruption buffer (100 mM Tris–HCl, pH 7.5, 150 mM NaCl, 1 mM phenylmethanesulfonylfluoride (PMSF), 5 mM $MgCl_2 \cdot 6H_2O$ and 0.1% (v/v) Triton X-100) was added, samples were vortexed and incubated for 30 minutes at 37 °C. To disrupt *E. coli* cells, samples were sonified on ice, using the Branson Model B-30 sonifier (Branson Sonic Power Co.) set at 60 pulses at 40% duty cycles at an output control of 2. To separate the membrane-bound insoluble and soluble protein fractions each of the samples was centrifuged at 21 500 x g for 15 minutes (4ºC). Supernatants (containing total soluble protein) were transferred into new microcentrifuge tubes and the pellets (containing insoluble fractions) were resuspended in 75 µl TBS. Samples were stored at 4 ºC.

4.2.6 Protein concentration determination and SDS-PAGE analysis

Protein concentrations of both the soluble and insoluble protein fractions were determined with the 2-D Quant kit (Amersham biosciences, USA). This kit is specifically designed to be compatible with reagents used in the preparation of protein samples, including detergents, reductants, chaotropes and carrier ampholytes. It is based on the principle of quantitatively precipitating proteins, while interfering substances remain in solution. Precipitated proteins are resuspended in a copper-containing solution and unbound copper is measured with colourimetric agent. The colour density is inversely related to the protein concentration.

From each protein sample a 2 µl aliquot was mixed with 100 µl precipitant and left to incubate at room temperature for 2-3 minutes. Co-precipitant (100 µl) was added, it was mixed by inversion and centrifuged at 10 000 x g for 5 minutes. Copper solution (20 µl) and double distilled de-ionised water (80 µl) was added and the samples were vortexed to dissolve the precipitated protein. For the colourimetric analysis 200 µl of working colour reagent (100 parts colour A with 1 part colour B) were added rapidly, to ensure instantaneous mixing. The total sample volume of 300 µl was transferred to a 96 well plate, incubated at room temperature for 15-20 minutes and the absorbance determined at 492 nm on a Multiscan Ascent scanner (Thermo Labsystems, Finland). Since the absorbance for this assay decreases with increasing protein concentration, an assay blank cannot be subtracted from all values. Therefore, water was used as reference. A standard curve was constructed using BSA.

SDS-PAGE was performed using a 4% stacking gel (0.5 mM Tris-HCl, 10% SDS, pH 6.8) and 10% separating gel (1.5 mM Tris-HCl, 10% SDS, pH 8.8). Acrylamide gels were prepared from acrylamide stock solution (30% Acrylamide, 2.67 % N'N'-Bis-methylene-acrylamide). Once prepared the gel solutions were polymerised with addition of 10% ammonium persulphate (75 µl to 15 ml separating gel and 30 µl to 7 ml stacking gel) and TEMED (N,N,N',N'-tetramethyl-ethylenediamine, 7.5 µl to 15 ml separating gel, 15 µl to 7 ml stacking gel). Samples were diluted 1:4 in reducing sample buffer (0.06 M Tris-HCl, 2% SDS, 0.1% glycerol, 0.05% β-mercaptoethanol, 0.025% bromophenol blue) and incubated at 37ºC for 10 minutes. Electrophoresis was performed in electrode running buffer (0.02 M Tris-HCl, 0.1 M glycine, 0.06% SDS, pH 8.3) using vertical BG-

verMINI gel electrophoresis systems (Baygene Biotech Company Limited, Hong Kong, China). All samples were run along with prestained ladders PageRuler™ (Fermentas GmbH, Germany) at an initial voltage of 80 V for 45 minutes followed by an increased voltage at 120 V for 4 hours. Gels were visualised after electrophoresis by Colloidal Coomasie Blue (CCB).

Colloidal Coomasie Brilliant Blue G250 stock solution (2% (v/v) phosphoric acid, 10% (w/v) ammoniumsulfate and 0.1% (v/v) Coomassie Brilliant Blue G250) was diluted with methanol (4:1 ratio) just before use. The gels were immersed in the Colloidal Coomassie solution and left shaking overnight. Gels were rinsed with a solution of 25% (v/v) methanol and 10% (v/v) acetic acid before destaining with 25% (v/v) methanol, until the background was clear (Neuhoff *et al.*, 1988). Gels were scanned on the GeneDoc™ XR+ System (Bio-Rad Inc., USA) and stored in 1% (v/v) acetic acid at 4ºC until submitted for mass spectrometry (MS).

4.2.7 Liquid Chromatography-Mass Spectrometry-Mass Spectrometry (LC-MS-MS) and protein sequence data analysis

SDS-PAGE gels were submitted to the Central Analytical Facility (CAF) at the University of Stellenbosch for LC-MS-MS, where all experiments were performed on a Thermo Scientific EASY-nLC II connected to a LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, Bremen, Germany) equipped with a nano-electrospray source.

Thermo Proteome Discoverer 1.3 (Thermo Scientific, Bremen, Germany) was used to identify proteins via automated database searching (Mascot, Matrix Science, London, UK and Sequest) of all tandem mass spectra against the UNIPROT (downloaded 27 July 2011) database. Carbamidomethyl cysteine was set as fixed modification, and oxidised methionine, N-acetylation and deamidation (NQ) was used as variable modifications. The precursor mass tolerance was set to 10 ppm, and fragment mass tolerance set to 0.8 Da. Two missed tryptic cleavages were allowed. Proteins were considered positively identified when they contained at least 2 tryptic peptides per proteins and a significant Mascot or Sequest score ($p<0.05$) as determined by Proteome Discoverer 1.3. Percolator was also used for validation of search results. In Percolator the NCBI (http//:ncbi.nlm.nih.gov), the Bmi gene index (BmiGI) of the DFCI

(http://compbio.dfci.harvard.edu/tgi/) and an assembled *R. microplus* contig database (assembled by Dr. C. Maritz-Olivier) were searched with a FDR (strict) of 0.02 and FDR (relaxed) of 0.05 with validation based on the q-value.

4.3 Results and Discussion

4.3.1 PCR amplification and cloning of full length BmMP1, BmMP2 and As51 into pGEM®-T Easy

A set of gene specific primers (GSPs) was designed for each of the 3 selected metzincins (BmMP1, BmMP2 and As51) (Table 4.4), allowing the amplification of the entire coding sequence depicted in the available sequence databases (see section 2.3 for full descriptive sequence details).

Table 4.4. Characteristics of the gene specific primers used for PCR amplification of BmMP1, BmMP2 and As51.

Primer Name	Primer sequence (5' to 3')	Tm (°C)
As51_FwC	TTG TCG TAG ACC TCT GTC AAG	59.7
As51_RvC	CGA GTA CAT AGC TTC GTG ACC	62.3
BmMP1_FwC	AGC ACC TTC AGC CAA AGA	60.4
BmMP1_RvC	GCA CAC TTC ACT TCT ATT GCG	63.3
BmMP2_FwC	GAA TCC TGG CAT CTT CTG C	62.1
BmMP2_RvC	TGC TCG TAG TTT ATT CAT TGC	60.9

Single stranded cDNA synthesised from *R. microplus* mixed lifestages total RNA served as template, to amplify the astacin and 2 reprolysin *R. microplus* transcripts. Upon analysis of the products on a 2% Agarose/TAE/GelRed gel it was clear that for all 3 transcripts a single band at the expected size was obtained (Figure 4.3).

Each of the 3 metzincins transcripts were cloned into pGEM® -T Easy vector and transformed into JM109 *E. coli.* Recombinant clones were identified with PCR screening and recombinant plasmids were submitted for automated nucleic acid sequencing. For each transcript the obtained sequence was aligned with the original sequence (Figure 4.4 and 4.5).

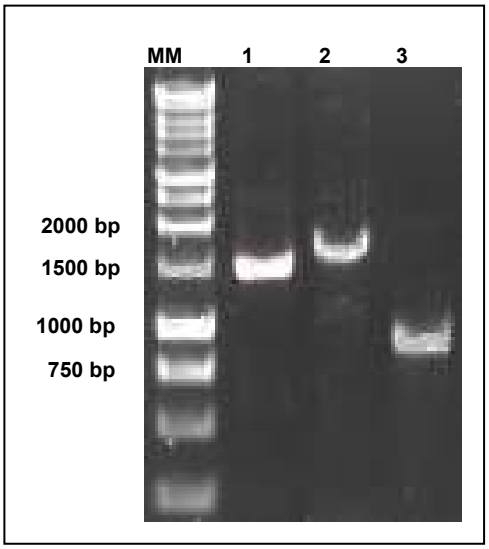

Figure 4.3. PCR amplification product of *R. microplus* BmMP1, BmMP2 and As51 using GSPs and cDNA created from *R. microplus* mixed lifestages. MM represents the 1 Kbp molecular marker. Lanes correspond to: (1) BmMP1, (2) BmMP2 and (3) As51.

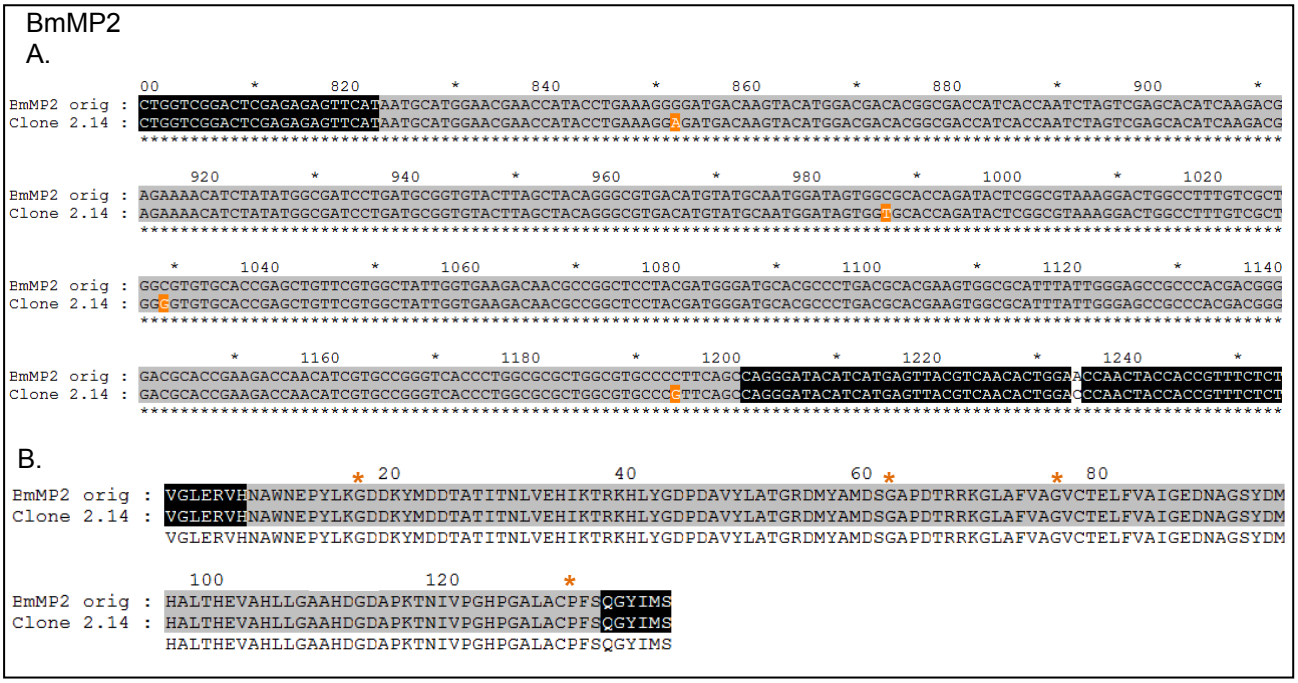

Figure 4.4. Nucleotide (A) and amino acid (B) sequence alignments of the BmMP2 fragment to be expressed compared to its original coding sequences. Grey reverse background indicates 100% similarity of the fragments to be expressed and orange single base pair mutations. Orange asterisks indicate synonomous mutations.

Sequencing results established that both BmMP2 and As51 were successfully amplified and cloned. For BmMP2, four synonomous single basepair mutations were observed.

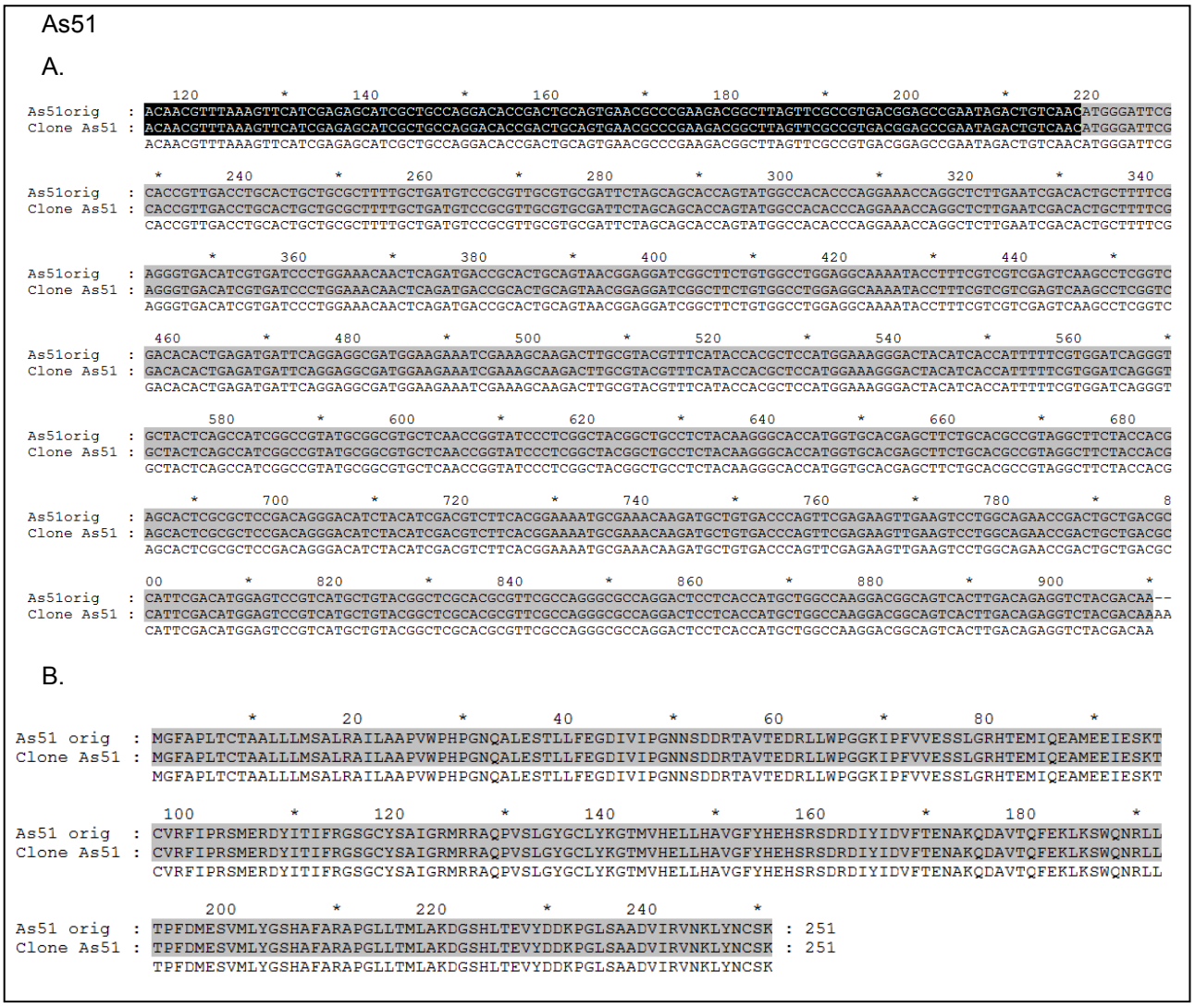

Figure 4.5. Nucleotide (A) and amino acid (B) sequence alignments of the As51 fragment to be expressed and compared to its original coding sequences. Grey reverse background indicates 100% similarity of the fragments to be expressed.

Therefore, for both BmMP2 and As51 the intact cloned constructs were ready for subsequent sub-cloning. However, after several optimisation strategies, (e.g. varying vector:insert ratio) no recombinant plasmid with intact BmMP1 could be obtained. To continue with the expression of BmMP1, a 453 bp synthetic gene construct was obtained from GenScript Inc. (USA).

4.3.2 Construction of pAFOR1x

The plasmid pAF0R1 was previously constructed by de la Fuente *et al.* (2003) to express a mutant MSP1α that lacks the tandem repeats. Briefly, the *msp1* gene (Oklahoma isolate clone per1; (de la Fuente *et al.*, 2001)) was amplified by PCR encoding a mutant MSP1α lacking 6 amino acids preceding the tandem repeats but containing 10 amino acids before the first putative transmembrane region (Garcia-Garcia *et al.*, 2004). The primers introduced an ATG initiation codon and *Eco*RI and *Bgl*II restriction sites for cloning into the pFLAG-CTC expression vector (Sigma, St. Louis, MO, USA) for expression in *E. coli* (Figure 4.6).

We received the pAFOR1-Bm95 (pMBAX) chimera construct (previously used by our collaborators at the IREC institute, University of Castilla-La Mancha, Spain). This chimera was constructed by PCR to encode a protein containing three peptides corresponding to BM95 amino acids 21–35, 132–147 and 397–410, respectively (Genbank accession number AAD38381) (Figure 4.6). Hence, to utilise the pAFOR1 vector in order to express BmMP1, BmMP2 and As51 selected domains chimerically with the modified MSP1, the obtained vector was reconstructed by removing BM95 and inserting additional basepairs to regain the function of the *Xho*I and *EcoR*I restriction sites (Figure 4.6).

Bm95 was successfully removed with enzyme digestion and the linearised plasmid (pFLAG-CTC-MSP1α/pAFOR1) purified. The modified linearised plasmid was amplified utilising gene specific primers containing the *Xho*I and *Eco*RI recognition sites (Figure 4.6), to obtain linearised plasmid (pAFOR1x) for subsequent blunt end ligation. Cell colonies present on the LB-agar/Ampicillin plates already suggested that ligation was successful, since intact plasmid is required for active ampicillin resistance. Plasmid isolation followed by single enzyme digestion and DNA gel electrophoresis confirmed that pAFOR1x was successfully reconstructed (Figure 4.6). From the gel analysis it was evident that both the *Xho*I and *Eco*RI cleavage sites were successfully reconstructed. Sequence analysis confirmed this and verified that all nucleotides were intact for in-frame cloning. Therefore, the pAFOR1x vector was readily available for chimeric expression of the 3 metzincins.

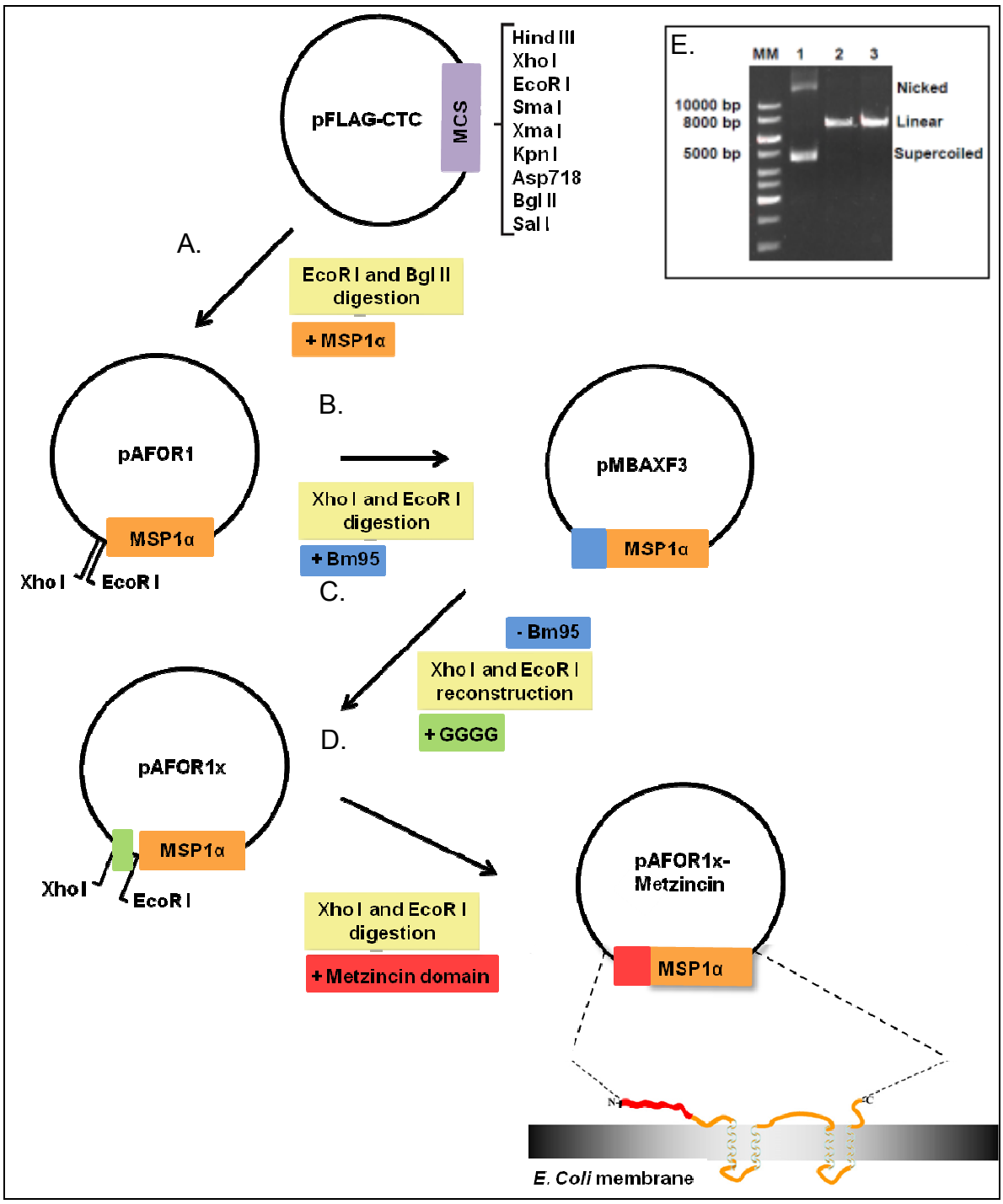

Figure 4.6. Schematic representation of the construction of the metzincin-MSP1α fusion surface expression vector. Mutant *A. mariginale* MSP1α (orange) (lacking the N-terminal repeats) was cloned into the pFLAG-CTC vector to construct pAFOR1 (A). Bm95 (blue) was fused with MSP1α to generate the Bm95-MSP1α expression vector pMBAXF3 (Canales *et al.*, 2008a) (B). To reconstruct the vector Bm95 was removed and with the use of gene specific primers (containing additional basepairs (green)) the native MPS1 vector was rebuild, known as pAFOR1x (C). The four selected metzincin domains (red) (BmMP1, BmMP2, As51V and As51A) were separately ligated into pAFOR1x to generate the pAFOR1x-metzincin fusion protein surface expression vector (D). The gel picture (E) indicates the restriction enzyme digests of expression plasmid pAFOR1x. (Lane 1: Undigested control; lane 2: *Eco*RI digested product; lane 3: *Xho*I digested product. MM represents the 1 Kbp molecular marker).

4.3.3 Domain selection

Several *in silico* methods are available for the prediction of epitopes and promising antigen domains. Where in the past vaccine development relied exclusively on biochemical and immunological determination of epitopes, new *in silico* predictive software and epitope and structure databases not only reduce time constraining laboratory experiments, but also provide a bottleneck approach for the identification and selection of antigens and their antigenic domains. For comprehensive epitope prediction both B- and T-cell epitope properties, together with physiochemical properties of the amino acids present within the epitope should be taken into consideration. These properties include surface accessibility, flexibility, charge, hydrophilicity and secondary structure (Yang and Yu, 2009).

For each of the two reprolysins the domain which showed the highest binding strength in at least five of the six prediction programs and the highest VaxiJen score was chosen for expression (refer to section 4.2.3). The segments selected for expression was solely based on the sequence dependent epitope predictions in combination with the highest sequence independent antigenicity score. For BmMP1 a 144 amino acid domain closer to the N-terminal, lacking both the zinc-binding histidine motif and methionine-turn, was selected. For BmMP 2 a 126 amino acid domain comprising the zinc-binding motif but not the methionine-turn was selected. From Figure 4.8 it can be observed that strong sequence dependent epitope rich areas are also available both upstream and downstream of the selected BmMP2 domain. These respective 150 amino acid segments scored lower based on sequence independent antigenicity (VaxiJen analysis) and therefore were excluded. For the astacin, it was decided to express the entire known sequence, as no ideal single area could be identified. As51 was divided into two equally sized domains, a N-terminal 111 amino acids domain (As51V) and a C-terminal 119 amino acids domain (As51A). All domains were selected in a manner to have the least interference possible with secondary structure elements.

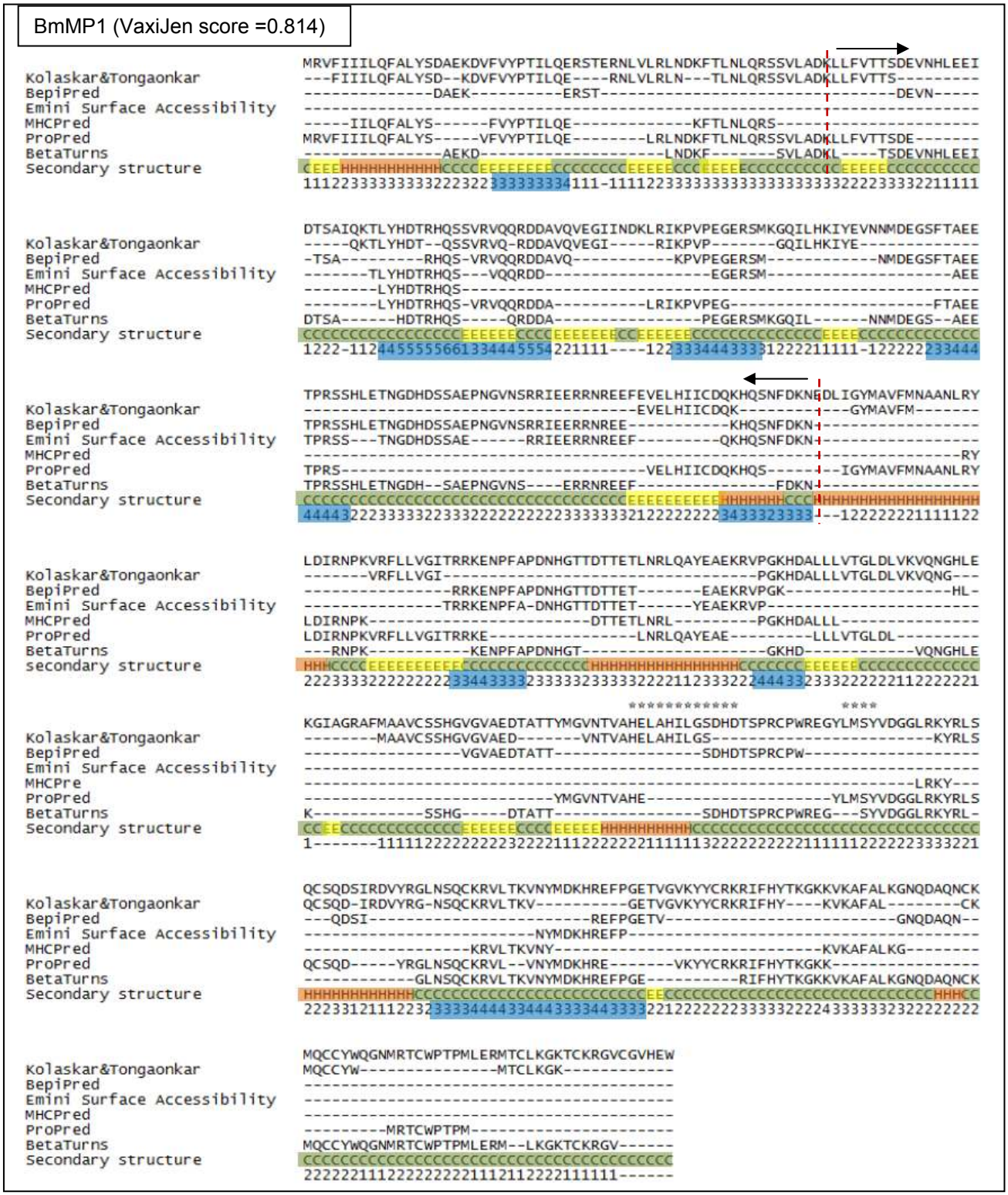

Figure 4.7. Results for prediction of epitopes and antigenic areas for BmMP1. Red dashed lines and arrows indicate the domain selected for expression. The highest predicted epitope regions are highlighted in blue and predicted secondary structure is denoted as H (orange), C (green) and E (yellow) representing α-helices, coils and beta-sheets, respectively. Asterisks (*) indicate the histidine motif and methionine turn.

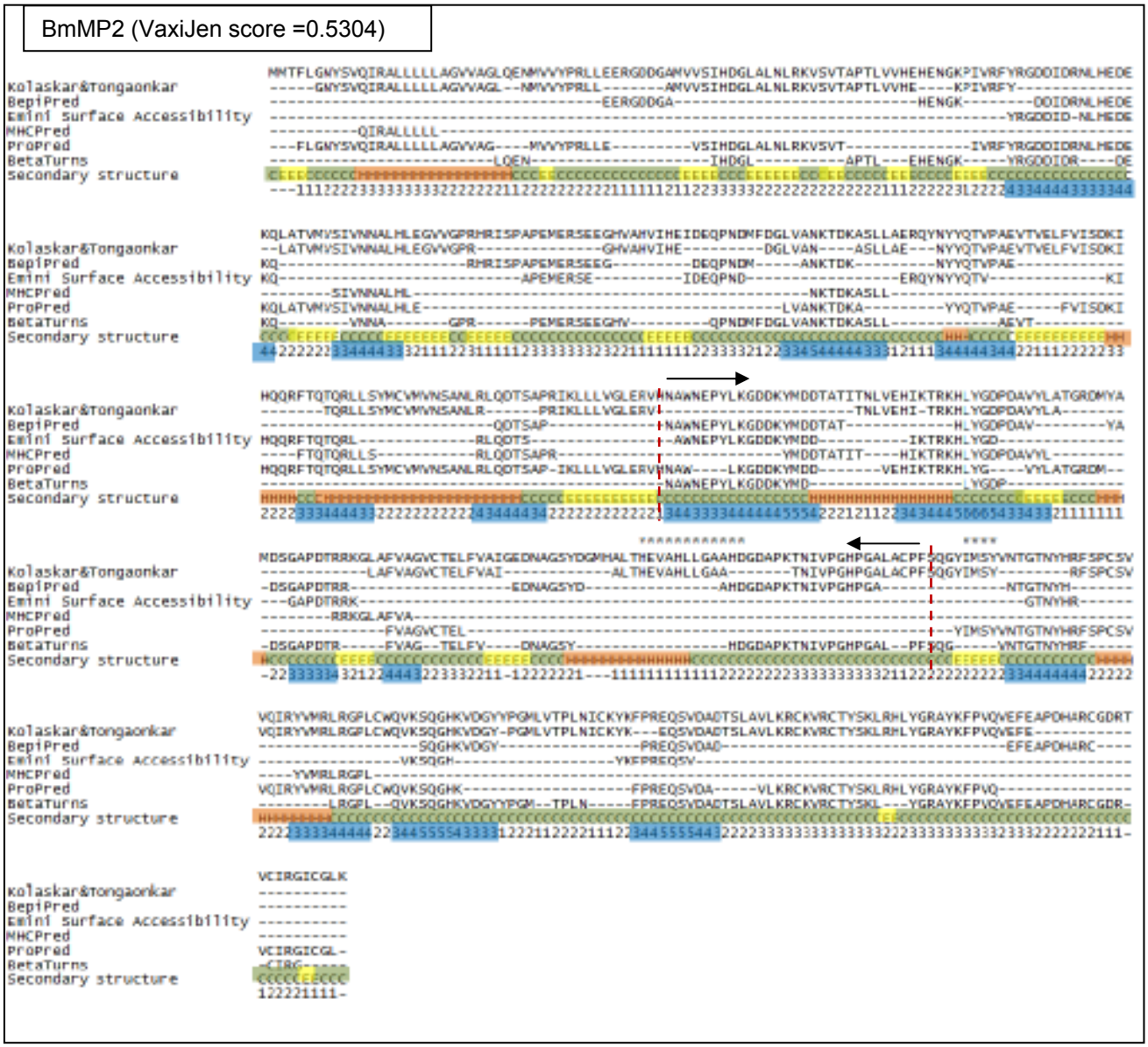

Figure 4.8. Results for prediction of epitopes and antigenic areas for BmMP2. Red dashed lines and arrows indicate the domain selected for expression. The highest predicted epitope regions are highlighted in blue and predicted secondary structure is denoted as H (orange), C (green) and E (yellow) representing α-helices, coils and beta-sheets, respectively. Asterisks (*) indicate the histidine motif and methionine turn.

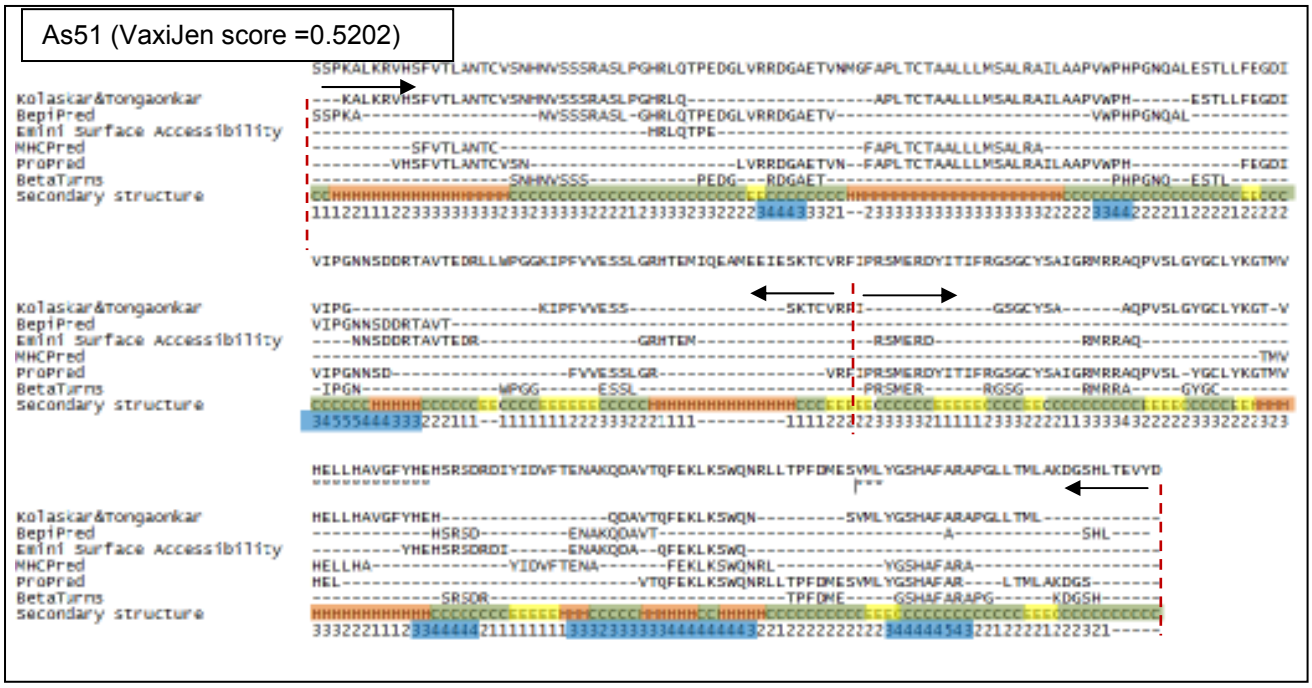

Figure 4.9. Results for prediction of epitopes and antigenic areas for As51. Red dashed lines and arrows indicate the domains selected for expression. The highest predicted epitope regions are highlighted in blue and predicted secondary structure is denoted as H (orange), C (green) and E (yellow) representing α-helices, coils and beta-sheets, respectively. Asterisks (*) indicate the histidine motif and methionine turn.

4.3.4 Directional sub-cloning into pAFOR1x

The selected metzincin domains for expression were successfully amplified and simultaneous digestion with *Xho*I and *Eco*RI restriction endonuclease was performed on all four amplified inserts and the purified pAFOR1x plasmid (Figure 4.10).

Table 4.5. Primers used in amplification and directional cloning of metzincin domains in pAFOR1x. Restriction enzyme sites are underlined. Entrokinase recognition site and the specific site of cleavage are highlighted in yellow and indicated by an arrow, respectively.

Primer Name	Primer sequence (5' → 3')	Tm (°C)
BmMP1_SC_Fw	CCGCTCGAGCAGAAAACTTTGTATCAT **EcoRI**	72.56
BmMP1_SC_Rv	CCGGAATTCCC↓TTTATCATCATCATCGTTTTTGTCAAAGTTACT **XhoI** **Entrokinase**	69.41
BmMP2_SC_Fw	CCGCTCGAGAATGCATGGAACGAACCATAC **EcoRI**	69.49
BmMP2_SC_Rv	CCGGAATTCCC↓TTTATCATCATCATCGCTGAAGGGGCACGCCAG **XhoI** **Entrokinase**	76.08
As51V_SC_Fw	CCGCTCGAGATGGGATTCGCACCGTTG **EcoRI**	71.04
As51V_SC_Rv	CCGGAATTCCC↓TTTATCATCATCATCGGTGATGTAGTCCCTTTCCA **XhoI** **Entrokinase**	73.87
As51A_SC_Fw	CCGCTCGAGATTTTTCGTGGATCAGGG **EcoRI**	68.01
As51A_SC_Rv	CCGGAATTCCC↓TTTATCATCATCATCGTCGTAGACCTCTGTCAA **XhoI** **Entrokinase**	73.14

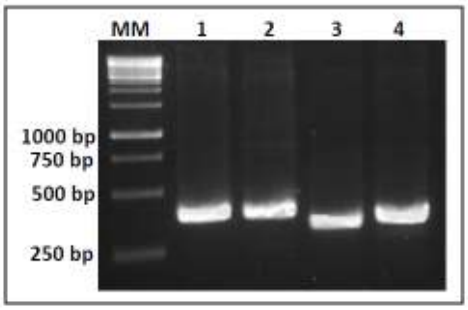

Figure 4.10. The PCR amplification products of the selected metzincin domains, using GSPs and the respective pGEM-metzincin plasmids as template. BmMP1 (lane 1), BmMP2 (lane 2), As51V (lane 3) and As51A (lane 4). MM represents the 1 Kbp molecular marker.

Figure 4.11-14 shows the nucleotide and amino acid sequences of two positive clones per metzincin domain, aligned with the expected sequences. All clones, showed 100% identity to the expected sequences. Therefore all metzicin-pAFOR1x constructs were ready for subsequent recombinant protein expression.

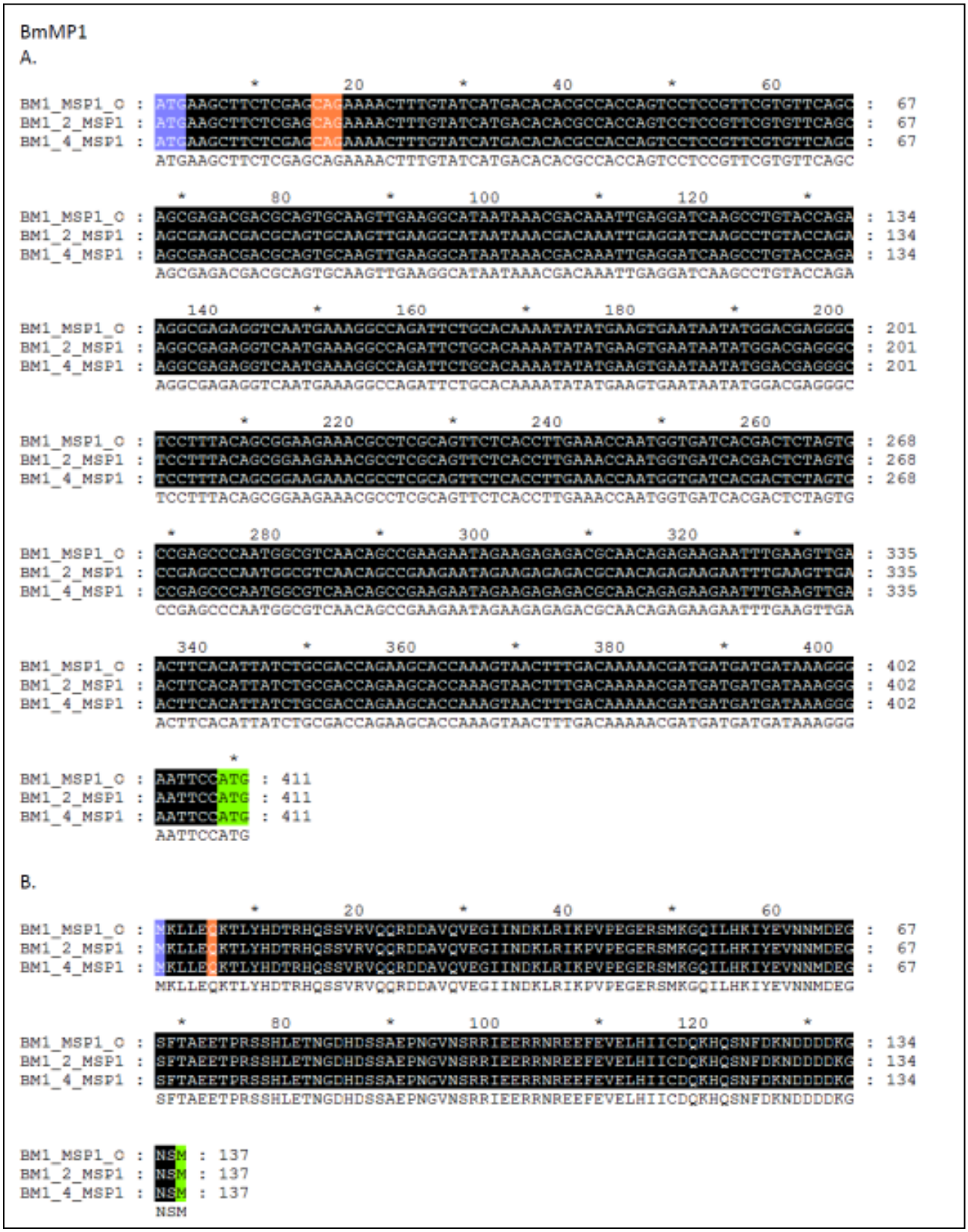

Figure 4.11. Nucleic acid (A) and amino acid (B) sequence alignments of two positive clones of the directionally cloned BmMP1 domain in the pAFOR1x vector. Black reverse background indicates 100% similarity to the expected sequence. Blue reverse background indicates the start codon of the expression (pFLAG-CTC) vector. The first amino acid codon for the BmMP1 domain and MSP1α is indicated by orange and green reverse background, respectively.

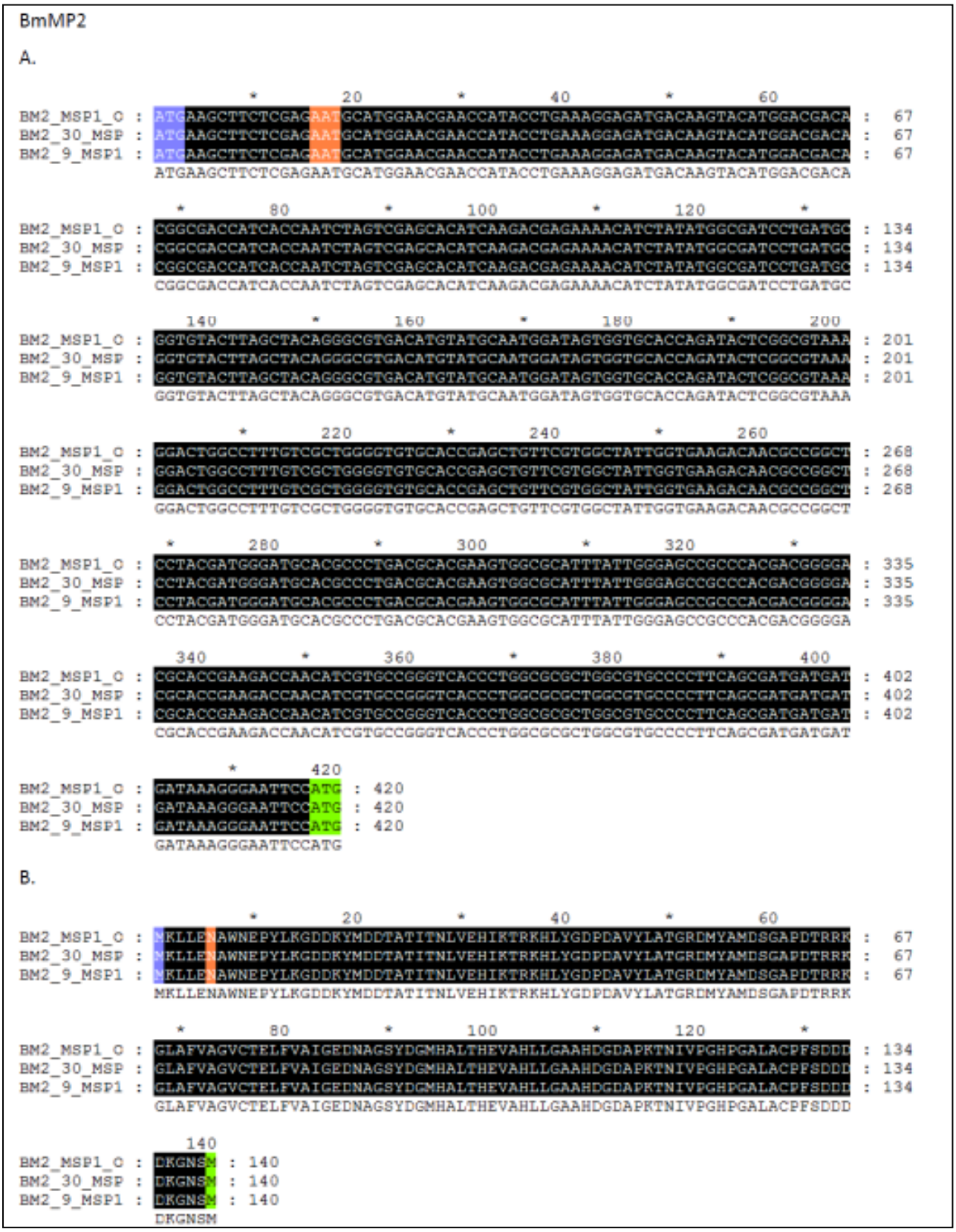

Figure 4.12. Nucleic acid (A) and amino acid (B) sequence alignments of two positive clones of the directionally cloned BmMP2 domain in the pAFOR1x vector. Black reverse background indicates 100% similarity to the expected sequence. Blue reverse background indicates the start codon of the expression (pFLAG-CTC) vector. The first amino acid codon for the BmMP2 domain and MSP1α is indicated by orange and green reverse background, respectively.

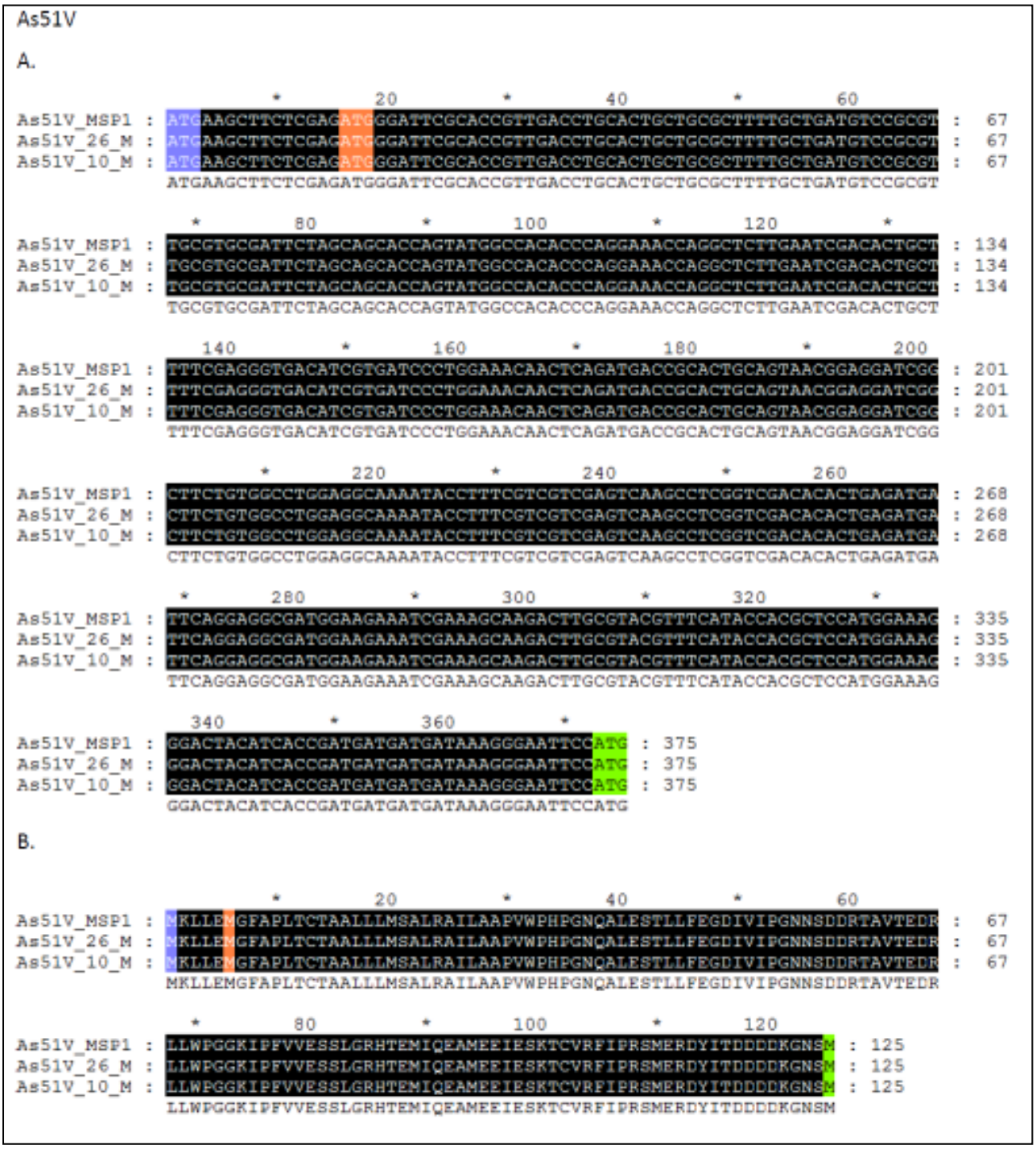

Figure 4.13. Nucleic acid (A) and amino acid (B) sequence alignments of two positive clones of the directionally cloned As51V domain in the pAFOR1x vector. Black reverse background indicates 100% similarity to the expected sequence. Blue reverse background indicates the start codon of the expression (pFLAG-CTC) vector. The first amino acid codon for the As51V domain and MSP1α is indicated by orange and green reverse background, respectively.

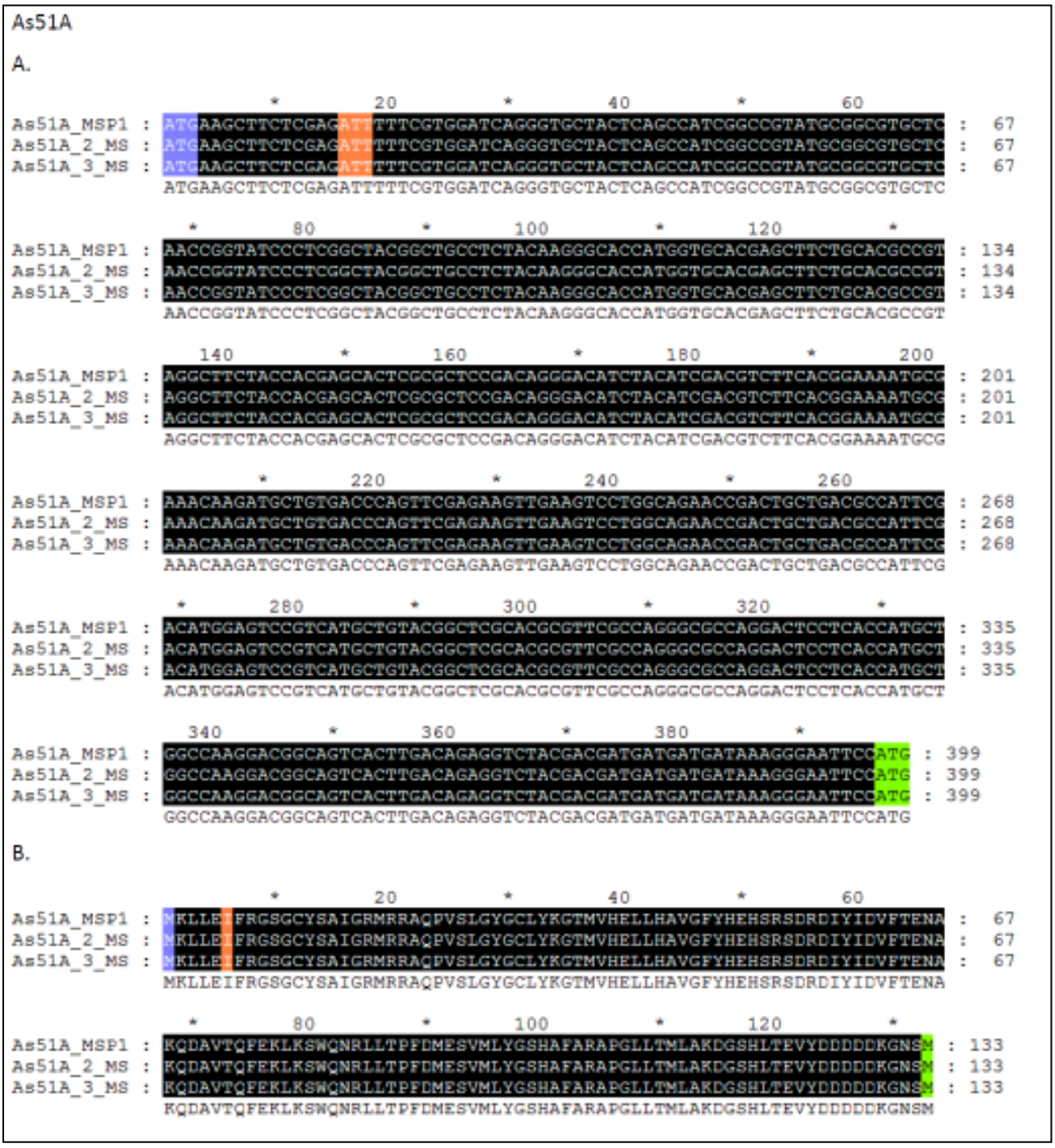

Figure 4.14. Nucleic acid (A) and amino acid (B) sequence alignments of two positive clones of the directionally cloned As51A domain in the pAFOR1x vector. Black reverse background indicates 100% similarity to the expected sequence. Blue reverse background indicates the start codon of the expression (pFLAG-CTC) vector. The first amino acid codon for the As51A domain and MSP1α is indicated by orange and green reverse background, respectively.

4.3.5 Expression of chimeric MSP1α-metzincins and LC-MS-MS confirmation

The final step for this study was to determine if the novel MSP1α - *E. coli* expression system can be utilised to express the 4 metzincin domains exoplasmic fused to MSP1α. Once the pAFOR1x-metzincin vectors were intact, expression was conducted with 2 clones of each construct. To establish optimal expression conditions expression was performed at three different temperatures (37°C, 28°C and 20°C) over a 12 hour time period. Figures 4.15 to 4.18 show the protein fractions of each clone at the different temperatures and time points on 10% SDS-PAGE gels. Native un-induced JM109 cells and pAFOR1x alone (expressing only MSP1α) was included as negative and positive controls, respectively.

Several optimisation strategies were initially conducted to obtain the optimal method for the analysis of the membrane fractions on the SDS-PAGE gels. This included assessing different staining methods and different sample preparation strategies. Colloidal Coomasie Brilliant Blue staining proved to be the optimal staining method, allowing the best analysis of the membrane fractions. As described in section 4.2.6 the samples were not subjected to boiling after being mixed with the SDS sample buffer, but were incubated at 37°C for 10 minutes prior to loading. Unlike soluble proteins, which need to be boiled before loading, membrane protein samples should not be boiled, since boiling often results in aggregation which in turn affect the migration (Sagne *et al.*, 1996; Rath *et al.*, 2009). Therefore to establish the optimal sample preparation method, boiled samples were compared with samples incubated at 37°C for 10 minutes, indicating that boiling membrane fractions adversely affected their migration.

From the results it is evident that the native pAFOR1x vector successfully expressed the positive control MSP1α, yielding a distinct band at the expected size of 54.5 kDa. By comparing the native JM109 membrane fraction (negative control) to all other samples, it is apparent that multiple *E. coli* insoluble (membrane bound) proteins were present throughout all the samples. Gel analysis of the membrane fractions (expressed from the different pAFOR1x-metzincin constructs) however, still clearly show that expression of chimeric BmMP1, BmMP2 and As51A was achieved.

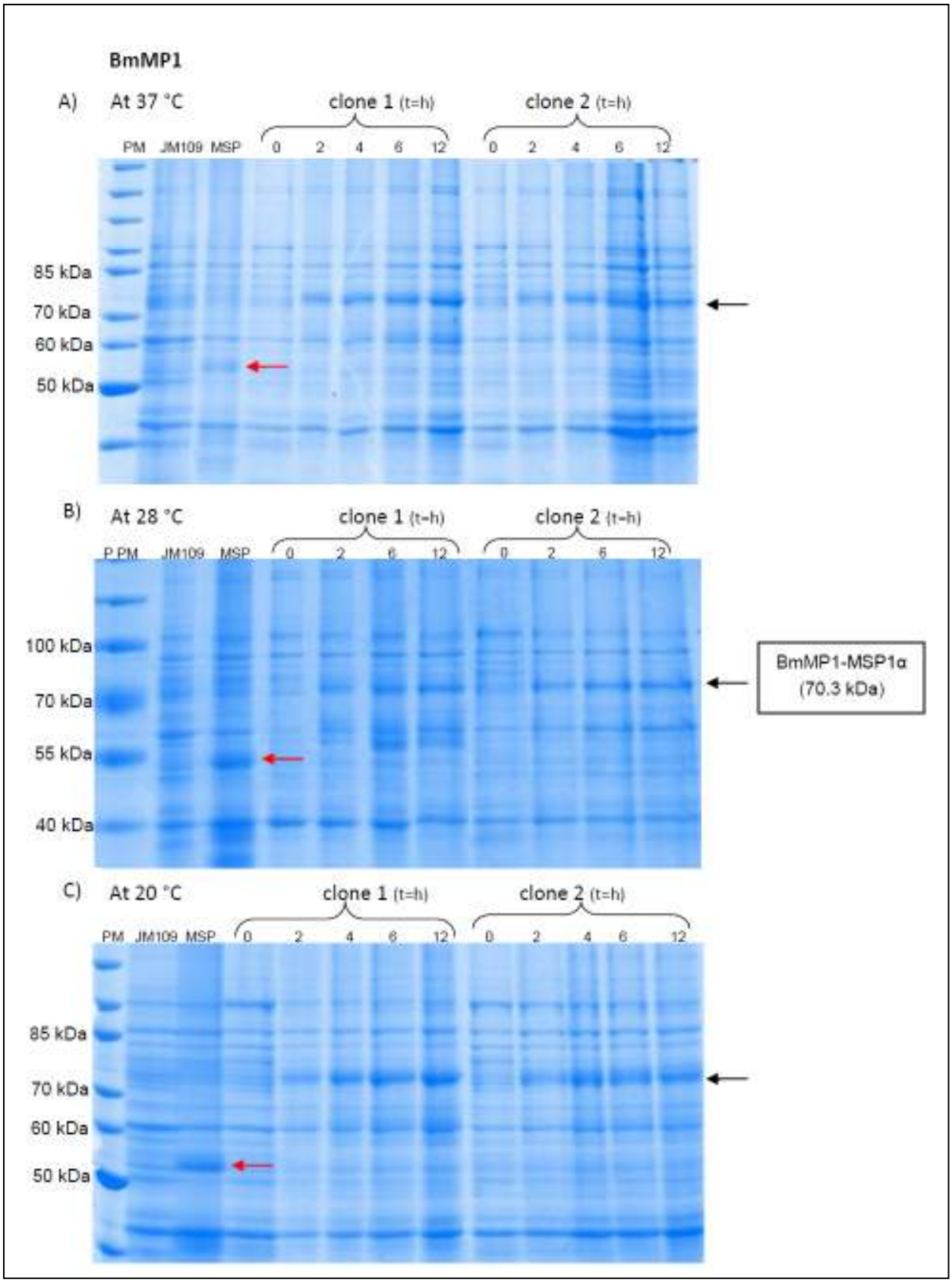

Figure 4.15. SDS-PAGE analysis of expression of BmMP1-MSP1α in JM109 *E. coli* cells. Native JM109 and MSP1α alone (54.5 kDa, indicated by red arrows) are shown in the first two lanes after the molecular protein marker (PM) or prestained marker (PPM). The different time points for each of the clones are shown in hours. The black arrows indicate the expected bands for BmMP1-MSP1α (70.3 kDa) at 37°C (A), 28°C (B) and 20°C (C).

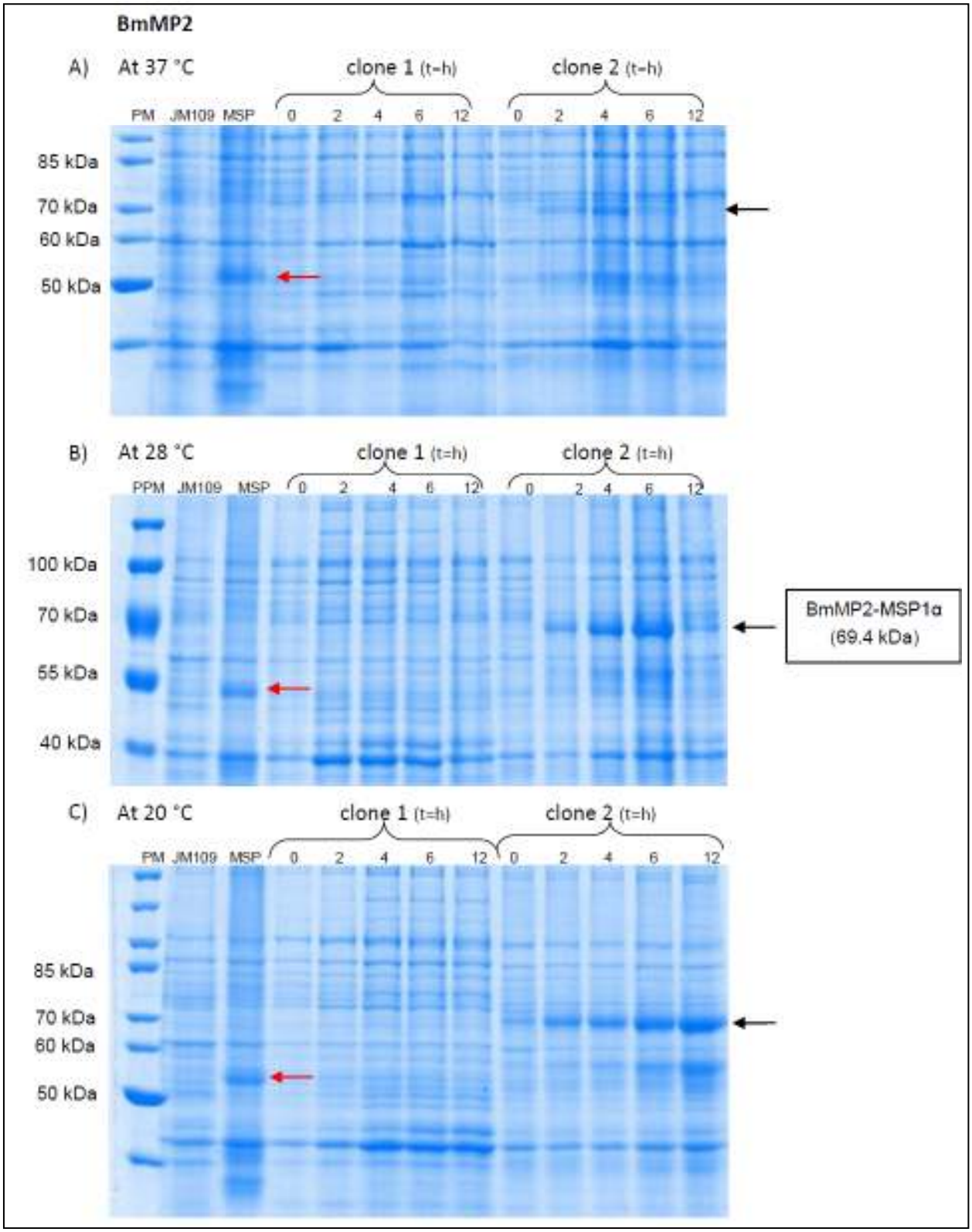

Figure 4.16. SDS-PAGE analysis of expression of BmMP2-MSP1α in JM109 *E. coli* cells. Native JM109 and MSP1α alone (54.5 kDa, indicated by red arrows) are shown in the first two lanes after the molecular protein marker (PM) or prestained marker (PPM). The time points for each of the clones are shown in hours. The black arrows indicate the expected band for BmMP2-MSP1α (69.4 kDa) at the 37°C (A), 28°C (B) and 20°C (C).

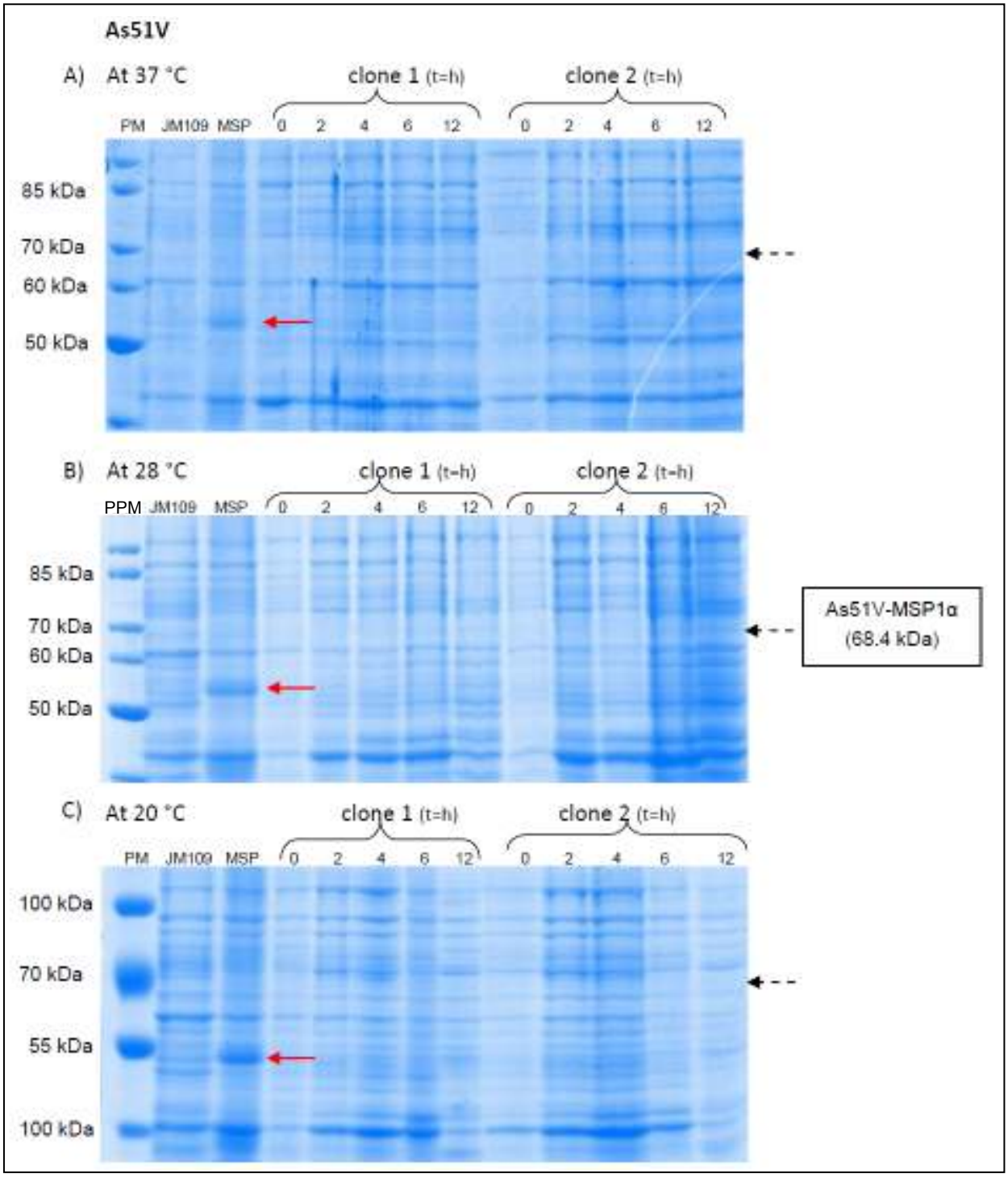

Figure 4.17. SDS-PAGE analysis of expression of As51V-MSP1α in JM109 *E. coli* cells. Native JM109 and MSP1α alone (54.5 kDa, indicated by red arrows) are shown in the first two lanes after the molecular protein marker (PM) or prestained marker (PPM). The time points for each of the clones are shown in hours. No bands were observed at the expected size for As51V-MSP1α (68.4 kDa, indicated by dashed arrows) at 37°C (A), 28°C (B) and 20°C (C).

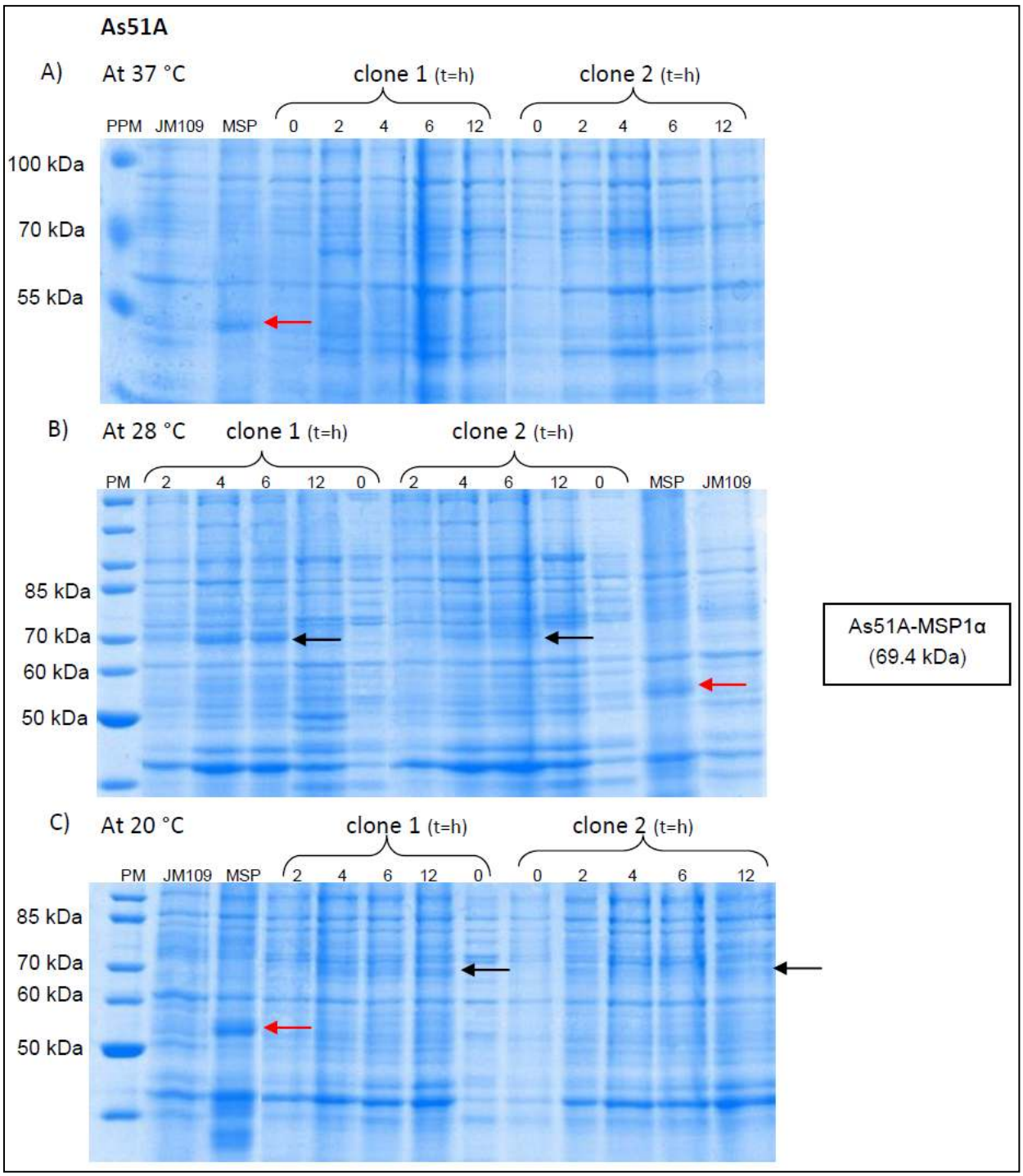

Figure 4.18. SDS-PAGE analysis of expression of As51A-MSP1α in JM109 *E. coli* cells. Native JM109 and MSP1α alone (54.5 kDa, indicated by red arrows) are shown in the first two lanes after the molecular protein marker (PM) or prestained marker (PPM) for A and C, and in the last two lanes for B. The time points for each of the clones are shown in hours. The black arrows indicate the expected band for BmMP2-MSP1α (69.4 kDa) at the 37°C (A), 28°C (B) and 20°C (C).

For BmMP1 the most distinct band, at the expected size of 70.3 kDa, was obtained for clone 1 at 20°C after 12 hours. For BmMP2 only clone 2 produced recombinant protein (69.4 kDa) at 28°C after 6 hours. For As51A, clone 1 produced a moderate intense band (at the expected size 69.4 kDa) at 28°C after 4 hours, but it was degraded and absent after 12 hours. This might have been due to the presence of proteases in the medium derived from cells that lysed overtime. LC-MS-MS was performed on the SDS-PAGE gels to validate the expression products. The respective metzincins sequences fused to MSP1α was confirmed, by good scoring peptides (Figure 4.19) and significant Mascot values (Table 4.6).

Table 4.6. LC-MS-MS analysis, for the confirmation of expressed domains. The abbreviation corresponds to (PI): Iso-electric point. * Indicate a significant p-value of smaller than 0.05 determined by Proteome Discoverer 1.3.

Domain	Mascot Score	% Coverage	# Unique peptides	Calculated PI
BmMP1	814.39*	37.89	14	5.68
BmMP2	12079.85*	54.94	8	5.52
As51A	10513.89*	62.33	9	6.24

Mascot results indicated fragment identity to the known expected sequences. The Mascot scores obtained for all 3 metzincin domains are well above the standard significant score of 70 (Perkins *et al.*, 1999), indicating that the recombinant expressed domains are in fact the respective metzincins.

For both As51V clones no recombinant chimeric protein, expected at 68.4 kDa, was obtained at any of the initial temperatures (37°C, 28°C or 20°C). First, a third confirmed in frame As51V-pAFOR1x clone (As51V_23) was expressed under the same initial conditions to account for any clone to clone variance that might have occurred. Again no recombinant chimeric As51V-MSP1α product was obtained. Based on previous findings, where it has been shown that invertebrate proteins which resist to express at ambient temperatures had success expressing at very low temperatures (Birkholtz *et al.*, 2008), all 3 As51V clones were submitted to expression at 4°C over 24 hours. However, gel analysis of the expressed membrane fractions revealed that As51V was

BmMP1-MSP1α

MKLLEQRTLY HDTRHQSSVR VQQRDDAVQV EGIINDKLRI KPVPEGERSM KGQILHKIYE VNNMDEGSFT AEETPRSSML
ETNGDHDSSA EPNGVNSRRI

EERRNREEFE VELHIICDQK HQSNFDKNDD DDKGNSMLAA NWRQEMRSKV ASVEYILAAR ALISVGVYAA QGEIAKSQGC
APLRVAEVEE IVRDGLVRSH

FHDSGLSLGS IRLVLMQVGD KLGLQGLKIG EGYATYLAQA FADNVVVAAD VQRGGACSAS LDSAIANVET SWSLHGGLVS
KDFDRDTKVE RGDLEAFVDF

MFGGVSYDDG NASAARSVLE TLAGHVDALG ISYNQLDKLD AGTLYSVVSF SADSAMDRGA VSDAADKFRV MMFGGAPAGQ
EKTAEPEHEA ATPSASSVLP

TVHGKVVDAV NRAKEAAEQA YAGVRKRYVA KPSDTTTQLV VAITALLITA FAICACLEPR LIGASGPLIW GCLALVALLP
LLGMAVHTAV SASSQKKAAG

GAQRVAAQER SRELSRARQE DQQKLHVPAI LTGLSVLVFI AAVVACIAVD ARRGTWQGSI CFLAAFVLFA ISAAVVMATR
DQSLAEECDS KCATARTAQA

VPGGQQQPRA TEGVVSGGSQ EGGAGVPGTS VPSAGSGSVP PATIMVSVDP QLVATLGAGV AQAAA

BmMP2-MSP1α

MKLLENAWNE PYLKGDDKYM DDTATITNLV EHIKTRKHLY GDPDAVYLAT GRDMYAMDSG APDTRRKGLA FVAGVCTELF
VAIGEDNAGS YDGMHALTHE

VAHLLGAAHD GDAPKTNIVP GHPGALACPF SDDDDKGNSM LAANWRQEMR SKVASVEYIL AARALISVGV YAAQGEIAKS
QGCAPLRVAE VEEIVRDGLV

RSHFHDSGLS LGSIRLVLMQ VGDKLGLQGL KIGEGYATYL AQAFADNVVV AADVQRGGAC SASLDSAIAN VETSWSLHGG
LVSKDFDRDT KVERGDLEAF

VDFMFGGVSY DDGNASAARS VLETLAGHVD ALGISYNQLD KLDAGTLYSV VSFSADSAMD RGAVSDAADK FRVMMFGGAP
AGQEKTAEPE HEAATPSASS

VLPTVHGKVV DAVNRAKEAA EQAYAGVRKR YVAKPSDTTT QLVVAITALL ITAFAICACL EPRLIGASGP LIWGCLALVA
LLPLLGMAVH TAVSASSQKK

AAGGAQRVAA QERSRELSRA RQEDQQKLHV PAILTGLSVL VFIAAVVACI AVDARRGTWQ GSICFLAAFV LFAISAAVVM
ATRDQSLAEE CDSKCATART

AQAVPGGQQQ PRATEGVVSG GSQEGGAGVP GTSVPSAGSG SVPPATIMVS VDPQLVATLG AGVAQAAA

As51A-MSP1α

MKLLEIFRGS GCYSAIGRMR RAQPVSLGYG CLYKGTMVHE LLHAVGFYHE HSRSDRDIYI DVFTENAKQD AVTQFEKLKS
RQNRLLTPFD MESVMLYGSH

AFARAPGLLT MLAKDGSHLT EVYDDDDDKG NSMLAANWRQ EMRSKVASVE YILAARALIS VGVYAAQGEI AKSQGCAPLR
VAEVEEIVRD GLVRSHFHDS

GLSLGSIRLV LMQVGDKLGL QGLKIGEGYA TYLAQAFADN VVVAADVQRG GACSASLDSA IANVETSWSL HGGLVSKDFD
RDTKVERGDL EAFVDFMFGG

VSYDDGNASA ARSVLETLAG HVDALGISYN QLDKLDAGTL YSVVSFSADS AMDRGAVSDA ADKFRVMMFG GAPAGQEKTA
EPEHEAATPS ASSVLPTVHG

KVVDAVNRAK EAAEQAYAGV RKRYVAKPSD TTTQLVVAIT ALLITAFAIC ACLEPRLIGA SGPLIWGCLA LVALLPLLGM
AVHTAVSASS QKKAAGGAQR

VAAQERSREL SRARQEDQQK LHVPAILTGL SVLVFIAAVV ACIAVDARRG TWQGSICFLA AFVLFAISAA VVMATRDQSL
AEECDSKCAT ARTAQAVPGG

QQQPRATEGV VSGGSQEGGA GVPGTSVPSA GSGSVPPATI MVSVDPQLVA TLGAGVAQAA A

Figure 4.19 Peptide fragments identified for BmMP1-, BmMP2- and As51A-MSP1α by LC-MS-MS analysis. Peptides that were detected by the LC-MS-MS for the expected specific protein are highlighted in colour. Red is indicative of low scoring peptides, yellow of moderate scoring peptides, and green of good scoring peptides. The initial methionine of MSP1α is indicated by a red arrow in each sequence. This is the provided output file from the sequencing facility, thus disrupted sequence have no significance.

still unable to be expressed. The exact explanation remains unknown. It is known that with multiple cell multiplication foreign DNA can mutate overtime. Single mutations can result in faulty translational initiation at ATG codons or addition or deletion of basepairs can even cause frameshifts, impeding the production of the wanted protein completely. Proteolytic enzymes, which are essential for endogenous protein processing may also lead to intracellular denaturation of foreign proteins after synthesis, or interfere with their correct assembly. Another possibility is that incorrect localisation of the chimeric protein might take place, the cell recognise it as foreign and degrades it. All the possible explanations for the lack of expression can cautiously be analysed to attempt successful expression of the As51V domain.

Fortunately the C-terminal domain As51A was successfully expressed, leaving one successful expressed domain for each of the *R. microplus* metzincins available for challenge against *R. microplus* infestation in cattle vaccine trails.

4.4 Conclusion

Vaccine formulations and developing production systems that will reduce the cost while increasing the immunogenicity of recombinant antigen is needed for proof-of-concept vaccine trails. Expression of full length recombinant metalloproteases present distinct challenges, due to their destructive nature and characteristics such as cysteine rich domains. Therefore, these proteins are usually expressed using complex systems such as mammalian cells, which are costly and yield low productivity. Hence, it was proposed to use a simple *E. coli* membrane expression system to express selected domains of the promising metzincin candidates (BmMP1, BmMP2 and As51) by fusing it to the *A. marginale* transmembrane protein, MSP1α. This system provides a new approach to establish the surface display of heterologous proteins on *E. coli* and suggests the possibility to use the recombinant bacteria for immunisation studies against cattle tick infestations, making it a simple vaccine production method for initial trials.

In this chapter, it was demonstrated that selected immunogenic domains of three *R. microplus* metzincins can be fused to the MSP1α N-terminal region and be expressed as recombinant protein displayed on the *E. coli* surface. First, gene specific primers which were designed to amplify the full length known sequence of each of the three *R. microplus* metzincins, successfully amplified each transcript and permitted successful A/T cloning into pGEM® T-easy. Secondly, to utilise the MSP1α based vector for the expression of four putative immunogenic metzincin domains, the received pMBAXF3 (MSP1α-Bm95) vector was reconstructed to pAFOR1x. Reconstruction entailed the removal of Bm95 and with the use of synthetic DNA primers the reconstruction and optimisation of the *Xho*I and *EcoR*I restriction enzyme sites.

For optimum expression the MSP1α expression system allows insertion of transcripts only up to 450 nucleic acids (150 amino acids). Thus, the next step was to select for each of the three metzincins the most immunogenic 150 amino acid domain. With the use of a variety of bio-informatic tools a three step strategy was followed. First, B- and T-cell epitope prediction programs were used to determine the most epitope rich region of each metzincin. For T cell epitope prediction MHC allele specificity predictions based on sequence are generally used. Conversely, B cell sequence based predictors remain very limited, since B cell epitopes are not in general continuous but rather

discontinuous. Prediction methods for continuous B cell epitopes are similar to T-cell epitope predictors, based on sequence and the amino acid properties. However in contrast, discontinuous B cell epitope prediction requires comprehensive insight of the three dimensional structure of the antigen-antibody complex. Thus, with the current lack of structures and bioinformatic tools for this purpose, discontinuous B cell epitope prediction remains very limited.

Secondly, it was established that once an epitope rich region was selected, that no important secondary structural features were disrupted. Finally, with the use of VaxiJen an alignment independent antigenicity score was calculated for each region, with the threshold cut-off value for being a probable antigen based on the parasite model set at 0.5. For BmMP1 and BmMP2, a 144 and a 126 epitope rich amino acid domain was selected, respectively. For As51 the transcript was divided into two near equal domains of 111 and 119 amino acids, respectively. All 4 domains were successfully directional cloned into the reconstructed pAFOR1x expression vector.

Both reprolysin domains were successfully expressed, producing recombinant BmMP1-MSP1α and BmMP2-MSP1α of 70.3 kDa and 69.4 kDa, respectively. The most optimum expression conditions for BmMP1 was at 20°C after 12 hours, whereas BmMP2 was best expressed at 28°C after 6 hours. The C-terminal domain of As51 was also successfully expressed fused to MSP1α at 28°C after 6 hours (with protein degradation observed at longer expression periods), leaving one successful expressed domain of each of the *R. microplus* metzincins available for cattle vaccine trails. It can be further investigated whether expression of BmMP1 at 20°C over a longer expression period can result in higher yields. However, longer expression periods might allow for increased proteolytic activity that may lead to product degradation and in effect lower yields rather than higher.

Even though the pioneers of the system use bioreactors, this study produced sufficient quantities with laboratory-scale expression. At optimum parameters, yields over 5 µg/µl were obtained for the insoluble membrane protein fraction, easily producing the adequate amount needed for vaccination (100 µg per injection). Although this expression system provides effective means to display immunogenic peptides to the host, it has several disadvantages.

Most of the limitations lie in using *E. coli* as the expression host. Although it delivers good quantities, it does not necessarily deliver good quality. *E. coli* is incapable of producing post-translation modification, such as glycosylation and disulphide bond formation that can be critical for the production of correctly folded protein, which in turn can drastically affect the protective capability of a vaccine candidate protein that relies critically on conformational epitopes. Therefore, the respective metzincin domains were selected in areas with limited cysteines and where simple secondary structure (i.e. coils with minimum α-helices and β-sheets) were predicted. Utilising the recombinant bacterial fractions as is, also takes us one step in reverse, back to “live-vaccines” from which most vaccine research have aimed to move away from. Exposure of live *E. coli* to the host may lead to or provoke several side effects.

To finally validate the protective ability of this combinatorial metzincin vaccine and determine the immunogenicity of the surface-displayed MSP1α-metzincin fusion proteins, it remains to be tested *in vivo*. The possible lack of recognition of the membrane-expressed protein, as it may be masked by other *E. coli* antigens recognised by cattle prior to the immunisation with MSP1α-metzincin, should be kept in mind. However, vaccination studies with membrane bacterial fractions containing MSP1α-BM95 chimeric protein have showed that a specific immune response was induced against the Bm95 antigen as well as the MPS1α, with the bacterial components adding an adjuvant effect. But, an entrokinase cleavage site has been incorporated to allow cleavage and isolation of the expressed metzincin domains if needed.

Immunisation of cattle with the recombinant *E. coli* membrane fraction containing the putative immunogenic metzincin domains holds the potential of being a cost effective and simple straightforward approach for the immunisation of cattle during first round trials. Once the proof of concept, for utilising a metzincin metalloprotease based anti-*R. microplus* vaccine has been demonstrated, time and knowledge will have to be invested to establish a more sustainable expression host and display vector combination.

4.5 References

Akashi, H. (2001). "Gene expression and molecular evolution." *Current Opinion in Genetics and Development* **11**(6): 660–666.

Almazán, C., O. Moreno-Cantú, J. A. Moreno-Cid, R. C. Galindo, M. Canales, M. Villar and J. de la Fuente (2012). "Control of tick infestations in cattle vaccinated with bacterial membranes containing surface-exposed tick protective antigens." *Vaccine* **30**(2): 265-272.

Anisuzzaman, M. K. Islam, M. A. Alim, T. Miyoshi, T. Hatta, K. Yamaji, Y. Matsumoto, K. Fujisaki and N. Tsuji (2012). "Longistatin is an unconventional serine protease and induces protective immunity against tick infestation." *Molecular and biochemical parasitology* **182**(1-2): 45-53.

Assenga, S. P., M. You, C. H. Shy, J. Yamagishi, T. Sakaguchi, J. Zhou, M. K. Kibe, X. Xuan and K. Fujisaki (2006). "The use of a recombinant baculovirus expressing a chitinase from the hard tick *Haemaphysalis longicornis* and its potential application as a bioacaricide for tick control." *Parasitology Research* **98**(2): 111-118.

Azhahianambi, P., J. de la Fuente, V. V. S. Suryanarayana and S. Ghosh (2009). "Cloning, expression and immunoprotective efficacy of rHaa86, the homologue of the Bm86 tick vaccine antigen, from *Hyalomma anatolicum anatolicum*." *Parasite Immunology* **31**(3): 111-122.

Balamurugan, V., G. R. Reddy and V. V. S. Suryanarayana (2007). "*Pichia pastoris*: A notable heterologous expression system for the production of foreign proteins - Vaccines." *Indian Journal of Biotechnology* **6**(2): 175-186.

Baneyx, F. (1999). "Recombinant protein expression in *Escherichia coli*." *Current opinion in biotechnology* **10**(5): 411-421.

Basaran, P. and E. Rodriguez-Cerezo (2008). "Plant molecular farming: Opportunities and challenges." *Critical Reviews in Biotechnology* **28**(3): 153–172.

Bensaci, M., D. Bhattacharya, R. Clark and L. T. Hu (2012). "Oral vaccination with vaccinia virus expressing the tick antigen subolesin inhibits tick feeding and transmission of *Borrelia burgdorferi*." *Vaccine* **30**(42): 6040–6046.

Birkholtz, L.-M., G. Blatch, T. L. Coetzer, H. C. Hoppe, E. Human, E. L. Morris, Z. Ngcete, L. Oldfield, R. Roth, A. Shonhai, L. Stephens and A. I. Louw (2008). "Heterologous expression of plasmodial proteins for structural studies and functional annotation." *Malaria Journal* **7**: 197.

Borghi, L. (2010). "Inducible gene expression systems for plants." *Methods in Molecular Biology* **655**: 65-75.

Brennan, F. R., T. D. Jones, L. B. Gilleland, T. Bellaby, F. Xu, P. C. North, A. Thompson, J. Staczek, T. Lin, J. E. Johnson, W. D. Hamilton and H. E. Gilleland, Jr. (1999). "*Pseudomonas aeruginosa* outer-membrane protein F epitopes are highly immunogenic in mice when expressed on a plant virus." *Microbiology* **145**(1): 211–220.

Bukau, B., E. Deuerling, C. Pfund and E. A. Craig (2000). "Getting newly synthesized proteins into shape." *Cell* **101**(2): 119-122.

Canales, M., C. Almazán, J. M. de la Lastra and J. de la Fuente (2008a). "*Anaplasma marginale* major surface protein 1a directs cell surface display of tick BM95 immunogenic peptides on *Escherichia coli*." *Journal of Biotechnology* **135**(4): 326-332.

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

Canales, M., J. M. de la Lastra, V. Naranjo, A. M. Nijhof, M. Hope, F. Jongejan and J. de la Fuente (2008b). "Expression of recombinant *Rhipicephalus* (*Boophilus*) *microplus*, *R. annulatus* and *R. decoloratus* Bm86 orthologs as secreted proteins in *Pichia pastoris*." *BMC Biotechnology* **8**: 14.

Carrió, M. M. and A. Villaverde (2002). "Construction and deconstruction of bacterial inclusion bodies." *Journal of Biotechnology* **96**(1): 3-12.

Casteleijn, M. G., A. Urtti and S. Sarkhel (2013). "Expression without boundaries: Cell-free protein synthesis in pharmaceutical research." *International Journal of Pharmaceutics* **440**(1): 39–47.

Cereghino, G. P. L., J. L. Cereghino, C. Ilgen and J. M. Cregg (2002). "Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*." *Current opinion in biotechnology* **13**(4): 329–332.

Cereghino, G. P. L. and J. M. Cregg (1999). "Applications of yeasts in biotechnology: Protein production and genetic analysis." *Current opinion in biotechnology* **10**(5): 422-427.

Cox, M. M. J. (2012). "Recombinant protein vaccines produced in insect cells." *Vaccine* **30**(10): 1759-1766.

da Silva Vaz, I., Jr., S. Imamura, K. Ohashi and M. Onuma (2004). "Cloning, expression and partial characterization of a *Haemaphysalis longicornis* and a *Rhipicephalus appendiculatus* glutathione S-transferase." *Insect Molecular Biology* **13**(3): 329–335.

de Boer, H. A. and R. A. Kastelein (1986). Biased codon usage: An exploration of its role in optimization of translation. *Maximizing Gene Expression*. W. Reznikoff and L. Gold. Boston, USA, Buttersworth.

de la Fuente, J., J. C. Garcia-Garcia, A. F. Barbet, E. F. Blouin and K. M. Kocan (2004). "Adhesion of outer membrane proteins containing tandem repeats of *Anaplasma* and *Ehrlichia* species (Rickettsiales: Anaplasmataceae) to tick cells." *Veterinary microbiology* **98**(3-4): 313-322.

de la Fuente, J., J. C. Garcia-Garcia, E. F. Blouin and K. M. Kocan (2003). "Characterization of the functional domain of major surface protein 1a involved in adhesion of the rickettsia *Anaplasma marginale* to host cells." *Veterinary microbiology* **91**(2-3): 265-283.

de la Fuente, J., J. C. Garcia-Garcia, E. F. Blouin, S. D. Rodríguez, M. A. García and K. M. Kocan (2001). "Evolution and function of tandem repeats in the major surface protein 1a of the ehrlichial pathogen *Anaplasma marginale*." *Animal Health Research Reviews* **2**(2): 163-174.

de la Fuente, J., K. M. Kocan, J. C. Garcia-Garcia, E. F. Blouin, P. L. Claypool and J. T. Saliki (2002). "Vaccination of cattle with *Anaplasma marginale* derived from tick cell culture and bovine erythrocytes followed by challenge-exposure with infected ticks." *Veterinary microbiology* **89**(2-3): 239-251.

Decrem, Y., M. Mariller, K. Lahaye, V. Blasioli, J. Beaufays, K. Z. Boudjeltia, M. Vanhaeverbeek, M. Cérutti, L. Vanhamme and E. Godfroid (2007). "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**(5): 549-560.

Desai, P. N., N. Shrivastava and H. Padh (2010). "Production of heterologous protein in plants: Strategies for optimal expression." *Biotechnology Advances* **28**(4): 427–435.

Domînquez, A., E. Ferminân, M. Sânchez, F. J. Gonzalez, F. M. Pérez-Campo, S. Garcia, A. B. Herrero, A. S. Vicente, J. Cabello, M. Prado, F. J. Iglesias, A. Choupina, F. J. Burguillo, L. Fernândez-Lago and M. Löpez (1998). "Non-conventional yeasts as hosts for heterologous protein production." *International Microbiology* **1**(2): 131-142.

Elvin, C. M. and D. H. Kemp (1994). "Generic approaches to obtaining efficacious antigens from vector arthropods." *International Journal for Parasitology* **24**(1): 67–79.

Emini, E. A., J. V. Hughes, D. S. Perlow and J. Boger (1985). "Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide." *Journal of Virology* **55**(3): 836-839.

Esposito, D. and D. K. Chatterjee (2006). "Enhancement of soluble protein expression through the use of fusion tags." *Current opinion in biotechnology* **17**(4): 353-358.

Fitchen, J., R. N. Beachy and M. B. Hein (1995). "Plant virus expressing hybrid coat protein with added murine epitope elicits autoantibody response." *Vaccine* **13**(12): 1051-1057.

Gao, J., J. Luo, R. Fan, V. Fingerle, G. Guan, Z. Liu, Y. Li, H. Zhao, M. Ma, J. Liu, A. Liu, Q. Ren, Z. Dang, C. Sugimoto and H. Yin (2008a). "Cloning and characterization of a cDNA clone encoding calreticulin from *Haemaphysalis qinghaiensis* (Acari: Ixodidae)." *Parasitology Research* **102**(4): 737-746.

Gao, J., J. Luo, R. Fan, G. Guan, V. Fingerle, C. Sugimoto, N. Inoue and H. Yin (2008b). "Cloning and characterization of a cDNA clone encoding troponin T from tick *Haemaphysalis qinghaiensis* (Acari: Ixodidae)." *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology* **151**(3): 323-329.

Garcia-Garcia, J. C., J. de la Fuente, K. M. Kocan, E. F. Blouin, T. Halbur, V. C. Onet and J. T. Saliki (2004). "Mapping of B-cell epitopes in the N-terminal repeated peptides of *Anaplasma marginale* major surface protein 1a and characterization of the humoral immune response of cattle immunized with recombinant and whole organism antigens." *Veterinary Immunology and Immunopathology* **98**(3-4): 137-151.

García-García, J. C., C. Montero, M. Rodríguez, A. Soto, M. Redondo, M. Valdés, L. Méndez and J. de la Fuente (1998). "Effect of particulation on the immunogenic and protective properties of the recombinant Bm86 antigen expressed in *Pichia pastoris*." *Vaccine* **16**(4): 374-380.

Gowda, D. C. and E. A. Davidson (1999). "Protein glycosylation in the malaria parasite." *Parasitology Today* **15**(4): 147-152.

Guan, P., I. A. Doytchinova, C. Zygouri and D. R. Flower (2003). "MHCPred: A server for quantitative prediction of peptide-MHC binding " *Nucleic Acids Research* **31**(13): 3621-3624.

Gustafsson, C., S. Govindarajan and J. Minshull (2004). "Codon bias and heterologous protein expression." *Trends in Biotechnology* **22**(7): 346–353.

Hall, T. A. (1999). "BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." *Nucleic Acids Symposium Series* **41**: 95-98.

Hartl, F. U. and M. Hayer-Hartl (2002). "Protein Folding. Molecular chaperones in the cytosol: From nascent chain to folded protein." *Science* **295**(5561): 1852-1858.

Hatta, T., K. Kazama, T. Miyoshi, R. Umemiya, M. Liao, N. Inoue, X. Xuan, N. Tsuji and K. Fujisaki (2006). "Identification and characterisation of a leucine aminopeptidase from the hard tick *Haemaphysalis longicornis*." *International Journal for Parasitology* **36**(10-11): 1123-1132.

Hericourt, F., S. Blanc, V. Redeker and I. Jupin (2000). "Evidence for phosphorylation and ubiquitinylation of the turnip yellow mosaic virus RNA-dependent RNA polymerase domain expressed in a baculovirus-insect cell system." *Biochemical Journal* **349**(2): 417-425.

Hodder, A. N., P. E. Crewther, M. L. S. M. Matthew, G. E. Reid, R. L. Moritz, R. J. Simpson and R. F. Anders (1996). "The disulfide bond structure of *Plasmodium* apical membrane antigen-1." *Journal of Biological Chemistry* **271**(46): 29446-29452.

Hunt, I. (2005). "From gene to protein: A review of new and enabling technologies for multi-parallel protein expression." *Protein Expression and Purification* **40**(1): 1–22.

Imamura, S., I. da Silva Vaz, Jr., S. Konnai, S. Yamada, C. Nakajima, M. Onuma and K. Ohashi (2009). "Effect of vaccination with a recombinant metalloprotease from *Haemaphysalis longicornis*." *Experimental and Applied Acarology* **48**(4): 345-358.

Imamura, S., I. da Silva Vaz, Jr., M. Sugino, K. Ohashi and M. Onuma (2005). "A serine protease inhibitor (serpin) from *Haemaphysalis longicornis* as an anti-tick vaccine." *Vaccine* **23**(10): 1301-1311.

James, D. C., R. B. Freedman, M. Hoare, O. W. Ogonah, B. C. Rooney, O. A. Larionov, V. N. Dobrovolsky, O. V. Lagutin and N. Jenkins (1995). "N-glycosylation of recombinant human interferon-ɣ produced in different animal expression systems." *Nature Biotechnology* **13**(6): 592-596.

Janin, J. and S. Wodak (1978). "Conformation of amino acid side-chains in proteins." *Journal of Molecular Biology* **125**(3): 357-386.

Kakker, N. K., M. V. Mikhailov, M. V. Nermut, A. Burny and P. Roy (1999). "Bovine leukemia virus Gag particle assembly in insect cells: Formation of chimeric particles by domain-switched leukemia/lentivirus Gag polyprotein." *Virology* **265**(2): 308-318.

Kalthoff, D., A. Giritch, K. Geisler, U. Bettmann, V. Klimyuk, H.-R. Hehnen, Y. Gleba and M. Beer (2010). "Immunization with plant-expressed hemagglutinin protects chickens from lethal highly pathogenic avian influenza virus H5N1 challenge infection." *Journal of Virology* **84**(22): 12002–12010.

Kaneda, M., S. M. Nomura, S. Ichinose, S. Kondo, K. Nakahama, K. Akiyoshi and I. Morita (2009). "Direct formation of proteo-liposomes by in vitro synthesis and cellular cytosolic delivery with connexin-expressing liposomes." *Biomaterials* **30**(23-24): 3971-3977.

Kaur, H. and G. P. S. Raghava (2004). "Prediction of α-turns in proteins using PSI-BLAST profiles and secondary structure information." *Proteins: Structure, Function and Genetics* **55**(1): 83-90.

Kolaskar, A. S. and P. C. Tongaonkar (1990). "A semi-empirical method for prediction of antigenic determinants on protein antigens." *Federation of European Biochemical Societies Letters* **276**(1-2): 172-174.

Kost, T. A., J. P. Condreay and D. L. Jarvis (2005). "Baculovirus as versatile vectors for protein expression in insect and mammalian cells." *Nature Biotechnology* **23**(5): 567-575.

Larsen, J. E., O. Lund and M. Nielsen (2006). "Improved method for predicting linear B-cell epitopes." *Immunome research* **2**: 2.

Lau, O. S. and S. S. M. Sun (2009). "Plant seeds as bioreactors for recombinant protein production." *Biotechnology Advances* **27**(6): 1015–1022.

LaVallie, E. R., E. A. DiBlasio, S. Kovacic, K. L. Grant, P. F. Schendel and J. M. McCoy (1993). "A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm." *Bio/Technology* **11**(2): 187-193.

Levitt, M. (1978). "Conformational preferences of amino acids in globular proteins." *Biochemistry* **17**(20): 4277-4285.

Lisowska, E. (2002). "The role of glycosylation in protein antigenic properties." *Cellular and Molecular Life Sciences* **59**(3): 445–455.

Liyou, N., S. Hamilton, C. Elvin and P. Willadsen (1999). "Cloning and expression of ecto 5'-nucleotidase from the cattle tick *Boophilus microplus*." *Insect Molecular Biology* **8**(2): 257-266.

Luckow, V. A., S. C. Lee, G. F. Barry and P. O. Olins (1993). "Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in *Escherichia coli*." *Journal of Virology* **67**(8): 4566-4579.

Luckow, V. A. and M. D. Summers (1988). "Signals important for high-level expression of foreign genes in *Autographa californica* nuclear polyhedrosis virus expression vectors." *Virology* **176**(1): 56-71.

Mason, H. S., J. M. Ball, J.-J. Shi, X. Jiang, M. K. Estes and C. J. Arntzen (1996). "Expression of Norwalk virus capsid protein in transgenic tobacco and potato and its oral immunogenicity in mice." *Proceedings of the National Academy of Sciences of the United States of America* **93**(11): 5335–5340.

Mena, J. A. and A. A. Kamen (2011). " Insect cell technology is a versatile and robust vaccine manufacturing platform." *Expert Review of Vaccines* **10**(7): 1063–1081.

Michel, E. and K. Wüthrich (2012). "High-yield *Escherichia coli*-based cell-free expression of human proteins." *Journal of biomolecular NMR* **53**(1): 43-51.

Miles, A. P., Y. Zhang, A. Saul and A. W. Stowers (2002). "Large-scale purification and characterization of malaria vaccine candidate antigen Pvs25H for use in clinical trials." *Protein Expression and Purification* **25**(1): 87-96.

Mulenga, A., C. Sugimoto, Y. Sako, K. Ohashi, A. Musoke, M. Shubash and M. Onuma (1999). "Molecular characterization of a *Haemaphysalis longicornis* tick salivary gland-associated 29-kilodalton protein and its effect as a vaccine against tick infestation in rabbits." *Infection and Immunity* **67**(4): 1652-1658.

Murphy, C. I., H. Piwnica-Worms, S. Grünwald, W. G. Romanow, N. Francis and H. Y. Fan (2004). Overview of the baculovirus expression system. *Current Protocols in Molecular Biology*. R. M. Asubel. Worcester, Massachusetts, USA, Aquila Biopharmaceuticals. **Chapter 16:** Unit 16 19.

Neuhoff, V., N. Arold, D. Taube and W. Ehrhardt (1988). "Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250." *Electrophoresis* **9**(6): 255-262.

Oberg, A. L., R. B. Kennedy, P. Li, I. G. Ovsyannikova and G. A. Poland (2011). "Systems biology approaches to new vaccine development." *Current Opinion in Immunology* **23**(3): 436-443.

Palomares, L. A., F. Kuri-Breña and O. T. Ramírez (2002). Industrial recombinant protein production. *The Encyclopedia of Life Support Systems*. Oxford, United Kingdom, EOLSS Publishers.

Patarroyo, J. H., R. W. Portela, R. O. de Castro, J. Couto Pimentel, F. Guzman, M. E. Patarroyo, M. I. Vargas, A. A. Prates and M. A. Dias Mendes (2002). "Immunization of cattle with synthetic peptides derived from the *Boophilus microplus* gut protein (Bm86)." *Veterinary Immunology and Immunopathology* **88**(3-4): 163-172.

Perkins, D. N., D. J. C. Pappin, D. M. Creasy and J. S. Cottrell (1999). "Probability-based protein identification by searching sequence databases using mass spectrometry data." *Electrophoresis* **20**(18): 3551-3567.

Porro, D. and D. Mattanovich (2004). Recombinant protein production in yeasts. *Recombinant Gene Expression: Reviews and Protocols*. P. Balbas and A. Lorence. Totowa, New Jersey, USA, Humana Press Inc. **267**.

Ramírez, O. T., G. K. Sureshkumar and R. Mutharasan (1990). "Bovine colostrum or milk as a serum substitute for the cultivation of a mouse hybridoma." *Biotechnology and Bioengineering* **35**(9): 882–889.

Rand, K. N., T. Moore, S. Sriskantha, K. Spring, R. Tellam, P. Willadsen and G. S. Cobon (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**(24): 9657-9661.

Rath, A., M. Glibowicka, V. G. Nadeau, G. Chen and C. M. Deber (2009). "Detergent binding explains anomalous SDS-PAGE migration of membrane proteins." *Proc.Natl Acad. Sci. USA* **106**(6): 1760-1765.

Renard, G., J. F. Garcia, F. C. Cardoso, M. F. Richter, J. A. Sakanari, L. S. Ozaki, C. Termignoni and A. Masuda (2000). "Cloning and functional expression of a *Boophilus microplus* cathepsin L-like enzyme." *Insect Biochemistry and Molecular Biology* **30**(11): 1017–1026.

Richardson, C. C. (1965). "Phosphorylation of nucleic acid by an enzyme from T4 bacteriophage-infected *Escherichia coli*." *Proceedings of the National Academy of Sciences of the United States of America* **54**(1): 158-165.

Richardson, M. A., D. R. Smith, D. H. Kemp and R. L. Tellam (1993). "Native and baculovirus-expressed forms of the immuno-protective protein BM86 from *Boophilus microplus* are anchored to the cell membrane by a glycosyl-phosphatidyl inositol linkage." *Insect Molecular Biology* **1**(3): 139-147.

Rigano, M. M., M. L. Alvarez, J. Pinkhasov, Y. Jin, F. Sala, C. J. Arntzen and A. M. Walmsley (2003). "Production of a fusion protein consisting of the enterotoxigenic *Escherichia coli* heat-labile toxin B subunit and a tuberculosis antigen in *Arabidopsis thaliana*." *Plant Cell Reports* **22**(7): 502–508.

Rigano, M. M. and A. M. Walmsley (2005). "Expression systems and developments in plant-made vaccines." *Immunology and Cell Biology* **83**(3): 271–277.

Rodriguez, M., R. Rubiera, M. Penichet, R. Montesino, J. Cremata, V. Falcon, G. Sanchez, R. Bringas, C. Cordoves, M. Valdes, R. Lleonart, L. Herrera and J. de la Fuente (1994). "High level expression of the *B. microplus* Bm86 antigen in the yeast *Pichia pastoris* forming highly immunogenic particles for cattle." *Journal of Biotechnology* **33**(2): 135-146.

Sagne, C., M. F. Isambert, J. P. Henry and B. Gasnier (1996). "SDS-resistant aggregation of membrane proteins: application to the purification of the vesicular monoamine transporter." *Biochemical Journal* **316 (Pt 3)**: 825-831.

Sahdev, S., S. K. Khattar and K. S. Saini (2008). "Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies." *Molecular and Cellular Biochemistry* **307**(1-2): 249-264.

Sharma, A. K. and M. K. Sharma (2009). "Plants as bioreactors: Recent developments and emerging opportunities." *Biotechnology Advances* **27**(6): 811-832.

Shi, Q., M. M. Lynch, M. Romero and J. M. Burns, Jr. (2007). "Enhanced protection against malaria by a chimeric merozoite surface protein vaccine." *Infection and Immunity* **75**(3): 1349-1358.

Shimizu, Y., Y. Kumura, B. W. Ying, S. Umekage and T. Ueda (2006). "Cell-free translation systems for protein engineering." *Federation of European Biochemical Societies Journal* **273**(18): 4133-4140.

Singh, B., M. Cabrera-Mora, J. Jiang and A. Moreno (2012). "A hybrid multistage protein vaccine induces protective immunity against murine malaria." *Infection and Immunity* **80**(4): 1491-1501.

Singh, H. and G. P. Raghava (2002). "ProPred: prediction of HLA-DR binding sites." *Bioinformatics* **17**(12): 1236-1237.

Sørensen, H. P. and K. K. Mortensen (2005). "Advanced genetic strategies for recombinant protein expression in *Escherichia coli*." *Journal of Biotechnology* **115**(2): 113-128.

Spirin, A. S., V. I. Baranov, L. A. Ryabova, S. Y. Ovodov and Y. B. Alakhov (1988). "A continuous cell-free translation system capable of producing polypeptides in high yield." *Science* **242**(4882): 1162-1164.

Streatfield, S. J. (2007). "Approaches to achieve high-level heterologous protein production in plants." *Plant Biotechnology Journal* **5**(2): 2-15.

Sturniolo, T., E. Bono, J. Ding, L. Raddrizzani, O. Tuereci, U. Sahin, M. Braxenthaler, F. Gallazzi, P. M. P., S. F. and H. J. (1999). "Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices." *Nature Biotechnology* **17**(6): 555-561.

Stutzer, C., B. J. Mans, A. R. M. Gaspar, A. W. H. Neitz and C. Maritz-Olivier (2009). "*Ornithodoros savignyi*: Soft tick apyrase belongs to the 5'-nucleotidase family." *Experimental Parasitology* **122**(4): 318-327.

Sugino, M., S. Imamura, A. Mulenga, M. Nakajima, A. Tsuda, K. Ohashi and M. Onuma (2003). "A serine proteinase inhibitor (serpin) from ixodid tick *Haemaphysalis longicornis*; cloning and preliminary assessment of its suitability as a candidate for a tick vaccine." *Vaccine* **21**(21-22): 2844-2851.

Tan, M., P. Huang, M. Xia, P.-A. Fang, W. Zhong, M. McNeal, C. Wei, W. Jiang and X. Jiang (2011). "Norovirus P particle, a novel platform for vaccine development and antibody production." *Journal of Virology* **85**(2): 753-764.

Tanner, W. and L. Lehle (1987). "Protein glycosylation in yeast." *Biochimica et Biophysica Acta - Reviews on Biomembranes* **906**(1): 81-99.

Tellam, R. D., D. L. Smith, D. H. Kemp and P. Willadsen (1992). Vaccination against ticks. *Animal parasite control utilising biotechnology*. W. K. Yong. Boca Raton, Florida, CRC Press.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins (1997). "The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools." *Nucleic Acids Research* **25**(24): 4876-4882.

Tinsley, T. W. and K. A. Harrap (1978). *Viruses of invertebrates*. New York, USA, Plenum.

Tolmasky, M. E., L. A. Actis and J. H. Crosa (1999). Plasmid DNA replication. *Encyclopedia of Bioprocess Techology*. New York, USA, John Wiley and Sons Inc.

Tremblay, R., D. Wang, A. M. Jevnikar and S. Ma (2010). "Tobacco, a highly efficient green bioreactor for production of therapeutic proteins." *Biotechnology Advances* **28**(2): 214–221.

Trimnell, A. R., R. S. Hails and P. A. Nuttall (2002). "Dual action ectoparasite vaccine targeting 'exposed' and 'concealed' antigens." *Vaccine* **20**(29-30): 3560-3568.

Trowitzsch, S., C. Bieniossek, Y. Nie, F. Garzoni and I. Berger (2010). "New baculovirus expression tools for recombinant protein complex production." *Journal of structural biology* **172**(1): 45-54.

Tsuda, A., A. Mulenga, C. Sugimoto, M. Nakajima, K. Ohashi and M. Onuma (2001). "cDNA cloning, characterization and vaccine effect analysis of *Haemaphysalis longicornis* tick saliva proteins." *Vaccine* **19**(30): 4287-4296.

Turnbull, I. F., D. R. Smith, J. Sharpp, S. Cobong and M. J. Hynes (1990). "Expression and secretion in *Aspergillus nidulans* and *Aspergillus niger* of a cell surface glycoprotein from the cattle tick, *Boophilus microplus*, by using the fungal amdS promoter system." *Applied and Environmental Microbiology* **56**(9): 2847-2852.

Vinarov, D. A., C. L. L. Newman and J. L. Markley (2006). "Wheat germ cell-free platform for eukaryotic protein production." *Federation of European Biochemical Societies Journal* **273**(18): 4160-4169.

Willadsen, P., R. V. McKenna and G. A. Riding (1988). "Isolation from the cattle tick, *Boophilus microplus*, of antigenic material capable of eliciting a protective immunological response in the bovine host." *International Journal for Parasitology* **18**(2): 183-189.

Willadsen, P., D. Smith, G. Cobon and R. V. McKenna (1996). "Comparative vaccination of cattle against *Boophilus microplus* with recombinant antigen Bm86 alone or in combination with recombinant Bm91." *Parasite Immunology* **18**(5): 241-246.

Wurm, F. and A. Bernard (1999). "Large-scale transient expression in mammalian cells for recombinant protein production." *Current opinion in biotechnology* **10**(2): 156–159.

Xu, J., X. Ge and M. C. Dolan (2011). "Towards high-yield production of pharmaceutical proteins with plant cell suspension cultures." *Biotechnology Advances* **29**(3): 278–299.

Yang, X. and X. Yu (2009). "An introduction to epitope prediction methods and software." *Reviews in Medical Virology* **19**(2): 77-96.

Yusibov, V., S. J. Streatfield and N. Kushnir (2011). "Clinical development of plant-produced recombinant pharmaceuticals: Vaccines, antibodies and beyond." *Human Vaccines* **7**(3): 313-321.

Zhou, J., H. Gong, Y. Zhou, X. Xuan and K. Fujisaki (2006a). "Identification of a glycine-rich protein from the tick *Rhipicephalus haemaphysaloides* and evaluation of its vaccine potential against tick feeding." *Parasitology Research* **100**(1): 77-84.

Zou, L., A. P. Miles, J. Wang and A. W. Stowers (2003). "Expression of malaria transmission-blocking vaccine antigen Pfs25 in *Pichia pastoris* for use in human clinical trials." *Vaccine* **21**(15): 1650-1657.

Chapter 5

Concluding Discussion

Problem statement, Rationale and Motivation

Although ticks are second to mosquitoes as global vectors of human diseases, they are arguably the most relevant vectors of disease-causing pathogens in domestic and wild animals. Factors such as climate, host movement, animal husbandry practices, vector distribution and vector population changes, affect tick distribution and occurrence of tick-borne diseases. The cattle tick, *Rhipicephalus microplus*, is regarded globally as the most economically important external parasite of cattle (Guerrero *et al.*, 2012). Infestations result in production losses for milk and beef as well as leather damage, but its greatest threat to the cattle industry is through its capability as a vector of the pathogens that cause anaplasmosis and babesiosis (Bock *et al.*, 2004; Merck & Co. Inc, 2008).This tick species has long been considered to be of minor importance in most of Africa, virtually absent throughout West Africa. However, in recent years it has been found that *R. microplus* has spread rapidly, infesting previously unaffected regions (Madder *et al.*, 2007). Moreover, it has been found that *R. microplus* has displaced “endemic” species, *R. decoloratus*, throughout much of its range in eastern and southern Africa, including the Limpopo province in South Africa (Tønnesen *et al.*, 2004). Therefore, given the fecundity of this species, its adaptability to different climatic zones, efficiency as a disease vector and ability to develop acaricide resistance, the full impact of its introduction to the African continent is difficult to estimate and likely to be catastrophic in the long term. The latter necessitates the development and implimentation of effective control strategies to alleviate the increasing pressure this species places on livestock.

Currently, control of *R. microplus* primarily relies on the use of chemical acaricides. Beside the increased number of *R. microplus* resistant populations, this method has several concerns regarding environmental, food and worker safety (Guerrero *et al.*, 2012). Furthermore, the escalating costs of acaricide development and marketing also factor into

what has developed as a real need for novel environmentally-sound control technologies. One of the most attractive alternative control approaches is to exploit the host's immunity (Willadsen, 2004). An anti-tick vaccine offers a promising alternative since vaccines, relative to chemical acaricides, are non-toxic, non-polluting, and less costly to develop and produce.

In complete ignorance of any microbes and immunology Edward Jenner pioneerd the first vaccine more than 200 years ago, which he named after vaccinia the cowpox virus (Lombard *et al.*, 2007). It was only 100 years later that Louis Pasteur developed the general principle governing vaccination. Vaccination operates by presenting a foreign antigen to the immune system of the host in order to evoke an immune response and to develop adaptive immunity to the infectious organism. Today, vaccination is generally considered to be the most successful application of immunological principles for human and animal health, globally preventing more than 6 million deaths and consequential savings of tens of billions $US annually (Ehreth, 2003). An effective vaccine, acts by inducing a specific long-lived immunity, being effective over a considerable period of time. To ascertain such characteristics several types of vaccines have been developed and succesfully applied, including: inactivated, live attenuated, subunit, virus-like particle, toxoid and DNA vaccines.

Although most vaccines generally target infectious viral or bacterial diseases, a limited number of vaccines against insect or arthropod vectors have been developed. In the late 1980's an effective vaccine against the cattle grub, *Hypoderma lineatum*, was developed and patented (Pruett et al., 1987; Pruett et al., 1989), but despite its efficacy, market factors prevented successful commercialization. The feasibility of vaccinating against *R. microplus*, has been demonstrated using the recombinant antigen Bm86 in commercially developed vaccines (de la Fuente *et al.*, 2007a). However the emergence of the variable efficacy of Bm86-based vaccines against different geographic *R. microplus* strains have encouraged research to deliver a next-generation vaccine that has an improved efficacy profile.

The most limiting step for the improvement and/or development of new vaccines remains the identification and characterization of effective vaccine antigen targets. Novel targets that are worth investigating for anti-tick vaccines are proteins which directly act in processes such as blood feeding, digestion, reproduction and development (e.g. proteases), in view of the fact that vaccination could inhibit these essential functions and ultimately affect tick survival (Guerrero *et al.*, 2012). An ideal target antigen would be a protein that is both essential for tick-vector survival and pathogen transmission, since vaccination will simultaneously control tick infestation and disease causing pathogen transmission. Based on various invertebrate and ixodid tick studies, metzincin metalloproteases have been recognized as key bioactive components in vital processes and pathogen transmission, offering a rationale for investigating these enzymes as potential tick vaccine targets.

Metalloproteases that form part of the array of known bioactive molecules in tick saliva have been classified as reprolysin-like metalloproteases based on sequence alignment analysis (Francischetti *et al.*, 2003). Based on sequence similarities, between the salivary gland metalloproteases and the haemorrhagic snake venom metalloproteases (SVMPs), and preliminary activity assays it is apparent that these metalloproteases contribute to the facilitation and maintenance of a fluid blood feeding cavity during extended feeding periods (Francischetti *et al.*, 2003). Accordingly, Decrem *et al.* (2007) proposed that these salivary gland metalloproteases are adequate vaccine candidates, since an elicited antibody response would result in the inhibition of essential physiological processes. Subsequent vaccination against these salivary gland metalloproteases did in fact affect the engorgement weight and fecundity of *I. ricinus* (found in Europe/ North Africa) (Decrem *et al.*, 2007) and *H. longicornis* (found in Japan) (Imamura *et al.*, 2009). These reports held up the notion that tick metzincin metalloproteases are promising anti-tick vaccine candidates, but required further investigation to fully explore their potential. Noting that these studies have only been performed for ticks affecting other parts of the world, such a study on South African strains of *R. microplus* will be of great value.

Prior to this study, metalloproteases have not been described in any other tick tissue than the salivary glands. However, various invertebrate studies have indicated metzincins (including reprolysins and astacins) in a wider variety of tissues with extensive roles in pathogen transmission (Gomis-Rüth, 2003), host infestation (Williamson *et al.*, 2006) and reproduction (Bowles *et al.*, 2008). Therefore, by means of a reverse genetics approach this study set out to identify metzincin homologues which exhibit similar vital functions in different *R. microplus* tick tissues, contributing to the tick's successful haematophagous parasitic behavior. Furthermore, we investigated the differential transcriptional response between the identified *R. microplus* metzincins in order to ascertain the development of a cocktail/ multi-metzincin vaccine, to lessen the compensation between the members of such a large protein family. The final aim was to utilize a simple *E. coli* based fusion protein membrane expression system to express selected epitope rich domains of the promising metzincins, in the most cost effective way for first round cattle trials.

Experimental overview and findings

With the use of several bio-informatic tools we commenced with the identification of astacin and reprolysin metalloproteases (MPs) present within the different available EST *R. microplus* databases **(Chapter 2)**. By means of BLAST searches, five coding sequences (BmMP1, BmMP2, BmMP3, BmMP4 and BmMP5) with significant similarity to a known reprolysin-like tick metalloprotease and three EST sequences (AS51, As70 and AsC) with significant similarity to the prototype astacin were identified. As with all members of the metzincins (Gomis-Rüth, 2009) the overall similarity and identity among the reprolysins and astacins were low. However, analysis of the amino acid sequences confirmed that the sequences contained their respective family's conserved zinc-binding his-motif as well as the conserved methionine, which clearly classified all eight *R. microplus* transcripts as metzincins. Phylogenetic analysis supports the alignment data and indicated that the 5 reprolysin and 3 astacin *R. microplus* homologues do belong to the reprolysin and astacin metzincin families, respectively. To elucidate function and localization the identified homologues were analysed for the presence of signal peptides, GPI-anchors and transmembrane regions. As expected all 5 full length *R. microplus* reprolysin-like

sequences contained a secretory signal peptide, indicating that these proteases are most likely secreted and act extracellularly. For the 3 *R. microplus* astacin-like MPs the entire coding sequences were not available and therefore the presence of signal peptides in the N-terminal ends of these metzincins could not be determined. Therefore, to obtain greater knowledge of the astacin-family's function within *R. microplus* an important future aim would be to obtain their full length coding sequences.

With the use of Reverse Transcriptase (RT)-PCR the expression profile of each *R. microplus* metzincin homologue was determined. Overall the results indicated that the reprolysins were most abundantly expressed in the salivary glands and ovaries, whereas the astacins were present in the midgut, salivary glands and reproductive organs. Therefore, taken all the results in consideration, we postulate that the 5 reprolysin-like metalloproteases are secreted by the salivary glands into the feeding cavity, where they act in maintaining a fluid intact feeding pool (Francischetti *et al.*, 2003). For the astacin-like metalloproteases it can be hypothesized, that within the midgut lumen these metzincins act on extracellular components to assure the blood does not clot within the midgut, prior to digestion. Since both As51 and AsC were detected at high levels in the reproductive organs we also hypothesize that these metzincins may be involved in the reproductive processes, oogenesis and spermatogenesis. As51 is the only transcript detected in *R. microplus* eggs and since astacins from other invertebrates are involved in embryogenesis or egg hatching (Bond and Beynon, 1995), it can be hypothesized that this *R. microplus* astacin may have similar function. However, the exact functions and substrates of these *R. microplus* metzincin metalloproteases are lacking due to lack of recombinant proteins and precise activity assays.

To evaluate the physiological importance of the eight metzincins within *R. microplus* we performed *in vivo* RNA interference. Double stranded RNA (targeting a randomly selected region within each transcript) was produced, and combinations of dsRNA were injected into freshly molted unfed females and engorged females to assess phenotypes of these in feeding female *R. microplus* and oviposited eggs **(Chapter 3)**. Upon evaluation of the data it was evident that the silencing of the two reprolysin-like transcripts, BmMP1 and BmMP2, and the astacin-like transcript, As51, yielded statistically significant phenotypes ($p \leq 0.05$),

with regards to the average egg weight and oviposition efficiency. Thus, this RNAi study has provided a tool to screen for tick-protective antigens in *R. microplus* and allowed for characterization of the effect and function of these metzincin. Taken together, the phenotypic results show support for the vital impact of the reprolysin and astacin metalloproteases within *R. microplus*.

The innovation to use RNAi as screening method for tick protective antigens was driven by the comparable results obtained between RNAi and Expression Library Immunisation (ELI) (de la Fuente *et al.*, 2005). However, RNAi as screening method has its limitations. The principle example is the absence of a statistical significant phenotype upon silencing of the highly effective antigen, Bm86 (Nijhof *et al.*, 2007). This directly demonstrates the possibility of missing potential candidates when utilizing RNAi as screening method. Other limitations to this system include off-target effects and fluctuations in gene expression during the tick's life cycle (de la Fuente *et al.*, 2007b; Nijhof *et al.*, 2007). The latter represent for example embryo and egg development genes. Silencing of the particular gene in freshly molted females is unlikely to result in silencing of the gene upon expression during egg development, producing deviant phenotypes. Despite these potential drawbacks, RNAi permits high throughput candidate evaluation and may be capable of identify vital transcripts, as in the case with promising candidate, subolesin (de la Fuente *et al.*, 2006). However, since RNAi is unable to resolve whether a target will be well presented to the host immune system and will provoke a protective immune response, a target can only be assertively verified by challenging the host and assessing the immune-biochemical effect.

RNAi serve as a valuable tool to study the gene expression and regulation of a network of essential genes, to reveal genes with redundant function and the effect of redundancy on metabolic pathways (Mohr and Perrimon, 2012). With the use of integrated semi-quantitative real-time PCRs and data analysis, this *in vivo* gene silencing RNAi study provided novel insight regarding gene regulation, between the different members of the metzincin clan investigated. It was found that several fold up-regulation of non-silenced transcripts occurred when other transcripts were specifically silenced. The remarkable finding was that this was a cross organ phenomena. Upon silencing of salivary gland

reprolysin metzincins the midgut and ovary astacin metzincins were activated and *vice versa*.

To date, all knowledge regarding metzincin gene regulation is restricted to what has been resolved for the matrix metalloportease (MMP) metzincin family. Because of MMPs active role in a diverse range of human diseases (including chronic inflammatory disorders and cancer), a lot of effort has been put into understanding the mechanisms of these metzincin genes (Rivera *et al.*, 2010). From various MMP studies it is evident that metzincin gene expression is regulated by differential signal transduction, in response to inductive stimuli (e.g. growth factor) (Vincenti and Brinckerhoff, 2007). However, very little to nothing is known about reprolysin and astacin gene regulation in invertebrates. Several tick studies have verified that most constituents of mechanisms pertaining to feeding and digestion are regulated at the transcription level (Franta *et al.*, 2010). However, as for the *R. microplus* astacins and reprolysins, the precise inductive stimuli and mechanisms triggering and regulating these tick proteins remain to be explored. Once characterized, these regulating factors in themselves may be promising candidates for tick vaccines. Of further interest is recent evidence of the significant contribution epigenetic modification has on metzincin regulation (Yan and Boyd, 2007). Chromatin remodeling, with histone acetylation and changes in methylation, is an important component in the induction and suppression of many genes.

Thus, it is clear that to fully explain the cross-organ phenomena observed between the *R. microplus* astacin and reprolysin metzincins, much more integrated studies on various levels are required. Ultimately, this can provide great insight into *R. microplus* gene regulation during feeding and digestion. For the moment, these current results support a combinatorial metzincin vaccine based on BmMP1, BmMP2 and As51, since such a metzincin-based vaccine might successful circumvent tick response against vaccination and therefore ultimately affect the reproductive fitness of *R. microplus*.

Considering the results and general characteristics of this clan of metalloproteases, metzincins also come with limitations when considered as vaccine candidates. Generally it

is believed that antigens concealed from the host (expressed in either gut or ovary tissues), are more attractive antigens than exposed antigens (secreted via the salivary glands), based on the hypothesis that co-evolution of the tick-host interaction has likely resulted in the tick's development of a means to circumvent the host's immune response to an exposed antigen (Brake and Perez de Leon, 2012). As secreted enzymes, the metzincins of the salivary glands (mainly the reprolysins) are exposed to the host's immune system during feeding. Furthermore, as proteases, metzincins are members of a very large gene family and a vaccine derived from a single member of such large family may fail for the reason of redundancy within the family, that allows for the function of the targeted family member to be taken up by other family members (Guerrero *et al.*, 2012). The latter is confirmed by the observation that silencing did not completely abolish metalloprotease activity. Thus the need for a combinatorial vaccine is stressed yet again.

Finally, we demonstrated that selected immunogenic domains of *R. microplus* metzincins can be fused to the MSP1α N-terminal region and be expressed as recombinant protein, displayed on the *E. coli* surface **(Chapter 4)**. This study has delivered one successful expressed domain for each of the three promising *R. microplus* metzincins. Although the MSP1α-based on-membrane expression system holds the advantage of being a cost effective and simple straightforward approach for the immunization of cattle in first round (proof of principle) trials, caution should be taken when considering the implementation of this system for commercial vaccine production. The strongest case in point is the possibility of lack of recognition of the target protein, as it may be masked by other *E. coli* antigens present in the bacterial membrane fraction. It should furthermore be taken into consideration that repeated exposure of *E. coli* to the host can lead to inflammation or disease, due to the presence of possible toxic substances. For that reason the non-pathogenic strain, JM109, is preferentially used. The answer however to alternative and improved systems is not simple.

A novel approach being investigated for tick control application is utilizing vaccinia virus (VV) as expression vector, where recombinant vector is administered to the host in an oral-based vaccine. This vaccination approach has been successfully used to eradicate rabies from foxes and raccoons in endemic areas in Canada utilizing VV in oral baits (Rosatte *et*

al., 2009). Bensaci *et al.* (2012) demonstrated the potential of this method for the control of tick infestation. The promising tick antigen, subolesin, was successfully expressed from VV and subsequent oral administration of VV-subolesin inhibited tick infestation with similar levels to that reported for recombinant subolesin administered percutaneously (Almazan *et al.*, 2005; Bensaci *et al.*, 2012). VV as vector for vaccine target delivery has the advantage of stably accommodating multiple gene insertions. In addition, oral immunization is a particularly attractive immunization method in developing countries as it is not invasive, requires less qualified administrators and is highly compatible with mass immunization. However, such oral-based vaccines have proven to be less immunogenic and the properties of the intestinal barrier could serve as an obstacle (Levine, 2010).

Future perspectives

The immense amount of sequencing data generated and the scientific and technological advances offered in the post genomic era, has resulted in a revolutionary paradigm shift in the field of vaccine development (Oberg *et al.*, 2011). Genome-based vaccine development (reverse vaccinology) now allow for the identification and characterization of all possible anti-tick targets, using available genome or transcriptome data in combination with computational analysis.

Maritz-Olivier *et al.* (2012) described a novel systematic approach that utilizes a combination of functional genomics (e.g. DNA microarray) techniques and a pipeline incorporating *in silico* prediction methods, permitting prediction of sub-cellular location and protective antigenicity, all in order to identify novel anti-tick vaccine targets. In the past, tick vaccine research has focused mainly on evaluating and producing extracted or recombinantly expressed protein that were selected on the basis of presumed vital function for the tick, cellular localization or immune recognition analyzed with expensive screening method that require a large number of test animals. However, with knowledge gained from the past it is now evident that in order to rationally asses a target, one should take several factors into consideration including: the presence of B- and T-cell epitopes in the target; the type of immune response to be elicited; the targets expression level as well as chemical

and physical properties of the antigen such as its post-translational modifications and aggregation status. With this new outlook on vaccine development, an accelerating growth is experienced in bioinformatic technologies and applications dealing with *in silico* analysis, modeling of immunological data and immunogenic predictions (altogether revered to as immunoinfomatics). However, one ultimate question that remains is how to prioritize a list of identified antigens for subsequent evaluation in expensive cattle trials. Evaluation methods that can be implemented include ranking putative candidates based on their ability to elicit a significant B cell response in smaller test animals such as rabbits and with the use of capillary feeding determine the phenotypical effect of the anti-sera on *R. microplus.*

The definite need for a more efficacious anti-tick vaccine, which ideally has to control multiple tick species in wide geographical areas, is now more apparent than ever. Like with all biological questions, not only one aspect can be taken into consideration. Crucial aspects that should also be taken into account when developing an anti-tick vaccine include: route of administration, adjuvants, use of model organisms, contaminating products from the expression system and the method to assess the efficacy. The final critical criteria for the next-generation anti-tick vaccine are that the necessary attention should be paid into taking antigen discovery into a viable commercial vaccine, focusing on correct marketing strategies and consumer acceptance; lest the discovery is to forever remain at the bench instead of providing the cattle industry with a much-needed anti-tick solution.

Importance of current study

To finally conclude: this study has made a definite contribution to the understanding of ixodid tick biology, specifically that of *R. microplus.* The most significant of these contributions (none of which has previously been reported in the literature) include:

i. Identification of astacin-like metzincins in ticks.
ii. Identification of metalloproteases in tick tissues other than the salivary glands. Based on the data of this study, it is evident that tick metzincins are not necessarily only

involved in extracellular matrix digestion (during blood feeding) but may also be involved in other processes such as oogenesis, spermatogenesis and egg hatching.

iii. We were able to validate the vital impact of the different *R. microplus* metzincins, by simultaneous dsRNA injection.

iv. We showed that due to protein redundancy, the effect of silencing on protein level cannot be observed with an overall metalloprotease assay.

v. With intergrated real-time PCR analysis we gained new insight in the understanding of gene regulation of ticks during feeding and digestion.

vi. Finally, we demonstrated that selected immunogenic domains of *R. microplus* metzincins can be fused to the MSP1α N-terminal region and be expressed as recombinant protein, displayed on the *E. coli* surface.

Utilising the knowledge gained from this study, the way forward will comprise cattle vaccination trials with recombinant immunogenic domains of the most promising metzincin candidates BmMP1, BmMP2 and As51 in a combinatorial vaccine. Should this provide protection future studies could include: 1) the investigation of the efficacy of the proposed metzincin vaccine against different tick species or geographical strains, 2) the exploration of an improved expression system, allowing essential translational modifications and 3) in depth transcriptome studies to obtain greater insight into the tick's compensation mechanism and potential resistance development.

References

Almazan, C., K. M. Kocan, E. F. Blouin and J. de la Fuente (2005). "Vaccination with recombinant tick antigens for the control of Ixodes scapularis adult infestations." *Vaccine* **23**(46-47): 5294-5298.

Bensaci, M., D. Bhattacharya, R. Clark and L. T. Hu (2012). "Oral vaccination with vaccinia virus expressing the tick antigen subolesin inhibits tick feeding and transmission of *Borrelia burgdorferi*." *Vaccine* **30**(42): 6040–6046.

Bock, R. E., L. A. Jackson, A. de Vos and W. K. Jorgensen (2004). "Babesiosis of cattle." *Parasitology* **129**: S247-S269.

Bond, J. S. and R. J. Beynon (1995). "The astacin family of metalloendopeptidases." *Protein Science* **4**(7): 1247-1261.

Bowles, V. M., A. R. Young and S. C. Barker (2008). "Mettalloproteases and egg-hatching in *Pediculus humanus*, the body (clothes) louse of humans (Phthiraptera: Insecta)." *Parasitology* **135**(1): 125-130.

Brake, D. K. and A. A. Perez de Leon (2012). "Immunoregulation of bovine macrophages by factors in the salivary glands of Rhipicephalus microplus." *Parasites & vectors* **5**: 38.

de la Fuente, J., C. Almazán, E. F. Blouin, V. Naranjo and K. M. Kocan (2005). "RNA interference screening in ticks for the identification of protective antigens." *Parasitology Research* **96**(3): 137-141.

de la Fuente, J., C. Almazán, E. F. Blouin, V. Naranjo and K. M. Kocan (2006). "Reduction of tick infections with *Anaplasma marginale* and *A. phagocytophilum* by targeting the tick protective antigen subolesin." *Parasitology Research* **100**(1): 85-91.

de la Fuente, J., C. Almazán, M. Canales, J. M. P. de la Lastra, K. M. Kocan and P. Willadsen (2007a). "A ten-year review of commercial vaccine performance for control of tick infestations on cattle." *Animal Health Research Reviews* **8**(1): 23-28.

de la Fuente, J., K. M. Kocan, C. Almazán and E. F. Blouin (2007b). "RNA interference for the study and genetic manipulation of ticks." *Trends in Parasitology* **23**(9): 427-433.

Decrem, Y., M. Mariller, K. Lahaye, V. Blasioli, J. Beaufays, K. Z. Boudjeltia, M. Vanhaeverbeek, M. Cérutti, L. Vanhamme and E. Godfroid (2007). "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**(5): 549-560.

Ehreth, J. (2003). "The global value of vaccination." *Vaccine* **21**(7-8): 596-600.

Francischetti, I. M. B., T. N. Mather and J. M. C. Ribeiro (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**(4): 869-875.

Franta, Z., H. Frantova, J. Konvickova, M. Horn, D. Sojka, M. Mares and P. Kopacek (2010). "Dynamics of digestive proteolytic system during blood feeding of the hard tick Ixodes ricinus." *Parasites & vectors* **3**: 119.

Gomis-Rüth, F.-X. (2003). "Structural aspects of the metzincin clan of metalloendopeptidases." *Molecular Biotechnology* **24**(2): 157-202.

Gomis-Rüth, F.-X. (2009). "Catalytic domain architecture of metzincin metalloproteases." *Journal of Biological Chemistry* **284**(23): 15353-15357.

Guerrero, F. D., R. J. Miller and A. A. Pérez de León (2012). "Cattle tick vaccines: Many candidate antigens, but will a commercially viable product emerge?" *International Journal for Parasitology* **42**(5): 421-427.

Imamura, S., I. da Silva Vaz, Jr., S. Konnai, S. Yamada, C. Nakajima, M. Onuma and K. Ohashi (2009). "Effect of vaccination with a recombinant metalloprotease from *Haemaphysalis longicornis*." *Experimental and Applied Acarology* **48**(4): 345-358.

Levine, M. M. (2010). "Immunogenicity and efficacy of oral vaccines in developing countries: lessons from a live cholera vaccine." *BMC biology* **8**: 129.

Lombard, M., P. P. Pastoret and A. M. Moulin (2007). "A brief history of vaccines and vaccination." *Rev Sci Tech* **26**(1): 29-48.

Madder, M., E. Thys, D. Geysen, C. Baudoux and I. Horak (2007). "*Boophilus microplus* ticks found in West Africa." *Experimental and Applied Acarology* **43**(3): 233-234.

Maritz-Olivier, C., W. van Zyl and C. Stutzer (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**(3): 179-187.

Merck & Co. Inc (2008). The Merck Veterinary Manual: Anaplasmosis. New Jersey, USA, Whitehouse Station.

Mohr, S. E. and N. Perrimon (2012). "RNAi screening: new approaches, understandings, and organisms." *Wiley interdisciplinary reviews. RNA* **3**(2): 145-158.

Nijhof, A. M., A. Taoufik, J. de la Fuente, K. M. Kocan, E. de Vries and F. Jongejan (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**(6): 653-662.

Oberg, A. L., R. B. Kennedy, P. Li, I. G. Ovsyannikova and G. A. Poland (2011). "Systems biology approaches to new vaccine development." *Current Opinion in Immunology* **23**(3): 436-443.

Pruett, J. H., C. C. Barrett and W. F. Fisher (1987). "Kinetic development of serum antibody to purified Hypoderma lineatum proteins in vaccinated and nonvaccinated cattle." *Southwestern Entomol* **12**: 79-88.

Pruett, J. H., J. G. Files, I. Kuhn and K. B. Temeyer (1989). Vaccines for the protection of animals against hypodermosis. European. **Patent 0326419**.

Rivera, S., M. Khrestchatisky, L. Kaczmarek, G. A. Rosenberg and D. M. Jaworski (2010). "Metzincin proteases and their inhibitors: foes or friends in nervous system physiology?" *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30**(46): 15337-15357.

Rosatte, R. C., D. Donovan, M. Allan, L. Bruce, T. Buchanan, K. Sobey, B. Stevenson, M. Gibson, T. MacDonald, M. Whalen, J. C. Davies, F. Muldoon and A. Wandeler (2009). "The control of raccoon rabies in Ontario Canada: proactive and reactive tactics, 1994–2007." *J Wildl Dis* **45**: 772–784.

Tønnesen, M. H., B. L. Penzhorn, N. R. Bryson, W. H. Stoltsz and T. Masibigiri (2004). "Displacement of *Boophilus decoloratus* by *Boophilus microplus* in the Soutpansberg region, Limpopo Province, South Africa." *Experimental and Applied Acarology* **32**(3): 199-208.

Vincenti, M. P. and C. E. Brinckerhoff (2007). "Signal transduction and cell-type specific regulation of matrix metalloproteinase gene expression: can MMPs be good for you?" *Journal of cellular physiology* **213**(2): 355-364.

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

Williamson, A. L., S. Lustigman, Y. Oksov, V. Deumic, J. Plieskatt, S. Mendez, B. Zhan, M. E. Bottazzi, P. J. Hotez and A. Loukas (2006). "*Ancylostoma caninum* MTP-1, an astacin-like mettaloprotease secreted by infective hookworm larvae, is involved in tissue migration." *Infection and Immunity* **74**(2): 961-967.

Yan, C. and D. D. Boyd (2007). "Regulation of matrix metalloproteinase gene expression." *Journal of cellular physiology* **211**(1): 19-26.