Identification and expression of proteases *C. sonorensis* and *C. imicola* important for African horsesickness virus replication

L Jansen van Vuuren

20272421

Dissertation submitted in partial fulfillment of the requirements for the degree *Magister Scientiae* in Biochemistry at the Potchefstroom Campus of the North-West University

Supervisor: Prof AA van Dijk

Co-supervisor: Prof TH Coetzer

May 2014

It all starts here ™

NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT

TABLE OF CONTENTS

AKNOWLEDGEMENTS

The success of this study was made possible by the contribution of several individuals. Without their input, effort and support this project would not have been possible.

I would like to express my sincere appreciation and thanks to the following people:

- My supervisor, Prof. Albie Van Dijk, for initiating this study, all her valuable insets, help and patience.
- My Co-supervisor, Prof. Theresa Coetzer, for all her help, support, trust and encouragements throughout this study.
- Dr. Gert Venter and his team at ARC-OVI for the collection and sorting of *C. imicola*.
- Prof. Christiaan Potgieter at Deltamune for his valuable discussions and providing me with the HS32/69 AHSV4 strain and BHK-21 cell line.
- Dr. Wouter van Wyngaardt for providing me with his original protocol for purifying AHSV and valuable discussions regarding AHSV purification.
- All the members of Prof. van Dijk's and Prof. Coetzer's laboratory for their support and assistance during this study.
- The National Research Foundation (NRF), Poliomyelitis Research Foundation (PRF) and the NWU for their generous financial support throughout this study.
- My parents, for giving me the opportunity to study this far and their on-going love and support.

ABBREVIATIONS

A

7-amino-4-methylcoumarin	AMC
African horsesickness	AHS
African horsesickness virus	AHSV
African horsesickness virus serotype 4	AHSV4
African horsesickness virus serotype 7	AHSV7
African horsesickness serotype 9	AHSV9
Afrika perdesiekte	APS
Afrika perdesiekte virus	APSV
Afrika perdesiekte serotpe 4	APSV4
Afrika perdesiekte serotpe 7	APSV7
Agricultural Research Council- Onderstepoort Veterinary Institute	ARC-OVI
Arginine	Arg
Avian Myeloblastosis Virus	AMV

B

Basepairs	bp
Basic local alignment software tool	BLAST
Benzoyl	BZ
Bluetongue/ Bloutong	BT
Bluetongue virus/ Bloutong virus	BTV
Benzoyl-L-arginine-AMC	BZ-L-Arg-AMC
Benzyloxycarbonyl-L-pyroglutamyl-glycyl-L-arginine-AMC	Z-Pyr-Gly-Arg-AMC
t-Butyloxycarbonyl-β-benzyl-L-aspartyl-L-prolyl-L-arginine-AMC	BOC-Asp(OBzl)-Pro-Arg-MCA
t-Butyloxycarbonyl-L-valyl-L-prolyl-L-arginine-MCA	BOC-Val-Pro-Arg-MCA

C

Carbobenzoxy-L-alanine-L-arginine-L-arginine-AMC	BOC-Ala-Arg-Arg-AMC
Caesium chloride	CsCl
Complimentary DNA	cDNA
Culicoides (Avaritia) imicola Kieffer	*C. imicola*
Culicoides bolitinos	*C. bolitinos*
Culicoides sonorensis	*C. sonorensis*
Culicoides sonorensis late trypsin	*CulsonLTRYP*
Cytopathic effect	CPE
Cytotoxic T-lymphocytes	CTL

D

Deoxyribonucleic acid	DNA
Dimethyl sulfoxide	DMSO
Dithiothreitol	DTT
Double-stranded RNA	dsRNA
Dulbecco's modified eagle medium	DMEM

E

Elution buffer	EB
Escherichia coli	*E. coli*
Ethelenediaminetetraacetic acid	EDTA

F

Foetal bovine serum	FBS

G

Glycine	Gly

I

Infectious sub-viral particles	ISVP
Integrated DNA technologies	IDT
International Committee on Taxonomy of Viruses	ICTV

Isopropyl β-D-1-thiogalactopyranoside	IPTG

L

Low Tris Buffer	LTB
Luria Broth	LB

N

National Center for Biotechnology Information	NCBI
Non-essential amino acids	NEAA
Non-structural	NS

O

Optical density	OD

P

Phosphate buffered saline	PBS
Polyacryamide gel electrophoresis	PAGE
Polymerase chain reaction	PCR
Polimerase ketting reaksie	PKR
Proline	Pro

R

Refractive index	RI
Reverse transcriptase	RT
Reverse transcriptase polymerase chain reaction	RT-PCR
Ribonucleic acid	RNA

S

Single stranded RNA	ssRNA
Sodium dodecyl sulfate	SDS
Sodium dodecyl sulfate polyacryamide gel electrophoresis	SDS-PAGE
Sodium hydroxide	NaOH
Super optimal broth with catabolite repression culture	SOC

T

Tetramethylethylenediamine	TEMED
Tris-Glycine SDS buffer	TGS Buffer
Tissue culture infectious dose at 50% assay	$TCID_{50}$
Tosyl phenylalanyl chloromethyl ketone	TPCK
Tris-acetate-EDTA	TAE

V

Viral protein/ Virale proteïen	VP

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

SUMMARY

African horsesickness (AHS) is one of the most deadly diseases of horses, with a mortality rate of over 90% in horses that have not been exposed to any African horsesickness virus (AHSV) serotype previously (Howell, 1960; Darpel *et al*., 2011). The Orbiviruses, African horsesickness virus (AHSV) and Bluetongue virus (BTV), are primarily transmitted to their mammalian hosts through certain haematophagous midge vectors (*Culicoides spp.*) (Erasmus, 1973). The selective cleavage of BTV and AHSV VP2 by trypsin-like serine proteases (Marchi *et al*., 1995) resulted in the generation of subsequent infectious sub-viral particles (ISVP) (Marchi *et al.,* 1995; van Dijk & Huismans, 1982). It is believed that this cleavage affects the ability of the virus to infect cells of the mammalian and vector host (Darpel *et al.*, 2011). Darpel *et al* (2011) identified a trypsin-like serine protease in the saliva of *Culicoides sonorensis* (*C. sonorensis*), which also cleaves the serotype determinant viral protein 2 (VP2) of BTV. And, a similar cleavage pattern was also observed by van Dijk & Huismans (1982) and Marchi *et al* (1995) with the use of trypsin and chymotrypsin. Manole *et al* (2012) recently determined the structure of a naturally occurring African horsesickness virus serotype 7 (AHSV7) strain with a truncated VP2. Upon further investigation, this strain was also shown to be more infective than the AHSV4 HS32/62 strain, since it outgrew AHSV4 in culture (Manole *et al*., 2012). Therefore, through proteolytic cleavage of these viral particles, the ability of the adult *Culicoides* to transmit the virus might be significantly increased (Dimmock, 1982; Darpel *et al*., 2011). Based on these findings, it is important to investigate the factors that influence the capability of arthropod-borne viruses to infect their insect vectors, mammalian hosts and their known reservoirs.

In this study, we postulated that one of the vectors for AHSV, *Culicoides imicola* (*C. imicola*), has a protease similar to the 29 kDa *C. sonorensis* trypsin-like serine protease identified by Darpel *et al* (2011). Proteins in the total homogenate of *C. imicola* were separated on SDS-PAGE and yielded several protein bands, one of which also had a molecular mass of around 29 kDa. Furthermore, proteolytic activity was observed on a gelatin-based sodium dodecyl sulfate polyacryamide gel electrophoresis (SDS-PAGE) gel. The activity of the protein of interest was also confirmed to be a trypsin-like serine protease with the use of class-specific protease inhibitors. A recombinant trypsin-like serine protease of *C. sonorensis* was generated using the pColdIII bacterial expression vector. The expressed protein was partially purified with nickel ion affinity chromatography. Zymography also confirmed proteolytic activity. With the use of the

protease substrates containing fluorescent tags and class specific protease inhibitors, the expressed protein was classified as a serine protease. It was also proposed that incubation of purified AHSV4 with the recombinant protease would result in the cleavage of AHSV4 VP2, resulting in similar VP2 digestion patterns as observed in BTV by Darpel *et al* (2011) or the truncated VP2 of AHSV7 by Manole *et al* (2012). BHK-21 cell cultured AHSV4 was partially purified through Caesium chloride gradient ultracentrifugation after which the virus was incubated with the recombinant protease. Since not enough virus sample was obtained, the outcome of VP2 digestion was undetermined.

In the last part of this study, it was postulated that *C. imicola* and *C. sonorensis* have the same trypsin-like serine protease responsible for the cleavage of VP2 based on the protease activity visualised in the whole midge homogenate. Since the genome of C. *imicola* is not yet sequenced, the sequence of this likely protease is still unknown. Therefore, we attempted to identify this *C. imicola* protease through polymerase chain reaction (PCR) amplification. Total isolated ribonucleic acid (RNA) of *C. imicola* was used to synthesize complementary deoxyribonucleic acid (cDNA). The cDNA was subjected to PCR using *C. sonorensis* trypsin-like serine protease-based primers. An 830 bp DNA fragment was amplified. However, sequence alignment and the basic local alignment software tool (BLAST), revealed that DNA did not encode with any other known proteins or proteases.

From the literature it seems that there is a correlation between the proteases in the vector and the mammalian species that succumb to AHS (Darpel *et al.,* 2011, Wilson *et al.,* 2009, Marchi *et al*., 1995). Based on the work performed in the study, a proteolytically active protein similar to the 29 kDa protein of *C. sonorensis* is present in *C. imicola*. The 29 kDa protease of *C. sonorensis* can also be expressed in bacteria which could aid in future investigations on how proteolytic viral modifications affect infectivity between different host species.

OPSOMMING

Afrika perdesiekte (APS) is een van die gevaarlikste siektes van perde, met 'n dodetal van bykans 90% in perde wat nooit voorheen aan enige serotipe van die virus blootgestel was nie (Howell, 1960; Darpel *et al*., 2011). Die Orbivirusse, Afrika perdesiekte virus (APSV) en Bloutong virus (BTV), word hoofsaaklik oorgedra na die soogdier gasheer deur middel van sekere bloed voedende muggie vektore (*Culicoides spp.*) (Erasmus, 1973). Die selektiewe kliewing van die BTV en APSV virus partikel deur tripsien-agtige serien proteases (Marchi *et al*., 1995) het tot die gevolg dat infektiewe sub-virale partikels gegenereer word (Marchi *et al.,* 1995; van Dijk & Huismans, 1982). Daar word ook beweer dat hierdie kliewing die vermoë van die virus om die insek en soogdier gasheer selle te infekteer beïnvloed (Darpel *et al*., 2011). Darpel *et al* (2011) het onlangs 'n tripsien-agtige serien protease in die speeksel van *C. imicola* ontdek wat die serotipe bepalende virale proteïen (VP) 2, van die BTV kleef. 'n Soortgelyke kliewings patroon op VP2 was ook waargeneem deur Van Dijk & Huismans (1982) en Marchi *et al* (1995) deur die behandeling van VP2 met tripsien en chimotripsien. Die struktuur van 'n nuwe APSV ras is onlangs deur Manole *et al* (2012) bepaal. Hierdie ras van die Afrika perdesiekte serotipe 7 (APSV7) besit 'n natuurlik gekleefde VP2. Na verdere ondersoek is daar vasgestel dat hierdie ras ook baie meer infektief is as die Afrika perdesiekte serotipe 4 (APSV4). Dit was bepaal deur groei kompetisie studies waar APSV7 baie vinniger as APSV4 in selkulture gegroei het. Daarvolgens kan daar vasgestel word dat die vermoë van die *Culicoides* vektor om die virus oor te dra merkwaardig verhoog word deur die kliewing van die buite dop proteïen, VP2 (Dimmock, 1982; Darpel *et al*., 2011). Daarom is dit belangrik om die faktore wat die vermoë van die virus beïnvloed om tussen die insek en soogdier gasheer oorgedra te word verder te ondersoek.

In die huidige studie het ons vermoed dat een van die vektore vir APSV, *C. imicola*, 'n protease besit wat soortgelyk is aan die reeds geïdentifiseerde tripsien-agtige serien protease van die BTV vektor, *C. sonorensis* (Darpel *et al*., 2011). Die proteïene in die totale muggie homogenaat was geskei op 'n SDS-PAGE gel en 'n verskeidenheid van proteïen bande was sigbaar. Een van hierdie proteïen bande was soortgelyke in grootte as die tripsien-agtige serien protease van *C. sonorensis*. Proteolitiese aktiwiteit van hierdie proteïen was bevestig deur 'n gelatien SDS-PAGE. Die gebruik van klas-spesifieke protease inhibeerders het vasgestel dat hierdie protease 'n tripsien-agtige protease is. 'n Rekombinante tripsien-agtige serien protease was saamgestel deur die nukleotied reeks van die *C. sonorensis* tripsien-agtige protease in die pColdIII bakteriële ekspressie vektor

in te kloneer. Die rekombinante protease was gesuiwer deur nikkel affiniteits ioon chromatografie. Proteolitiese aktiwiteit van die protease was deur ‘n reeks fluoreseerende protease substrate bepaal. Protease klas-spesifieke inhibeerders het die rekombinante protease geklassifiseer as 'n serien protease. Deur die inkubering van gesuiwerde BTV en die speeksel proteases in *C. sonorensis*, kon Darpel *et al* (2011) daarin slaag om die struktuur van BTV VP2 te verteer. Gebaseer hierop was APSV4 gekultiveer in die BHK-21 sellyn en gedeeltelik gesuiwer met behulp van CsCl digtheids gradiënt ultrasentrifugering. Die gedeeltelik gesuiwerde virus en rekombinante protease was saam geïnkubeer in verskillende tyds intervalle. Die virus opbrengs was wel so laag dat die effek van rekombinante protease behandeling op VP2 nie vasgestel kon word nie. Op grond van die *C. imicola* protease wat vroeër in hierdie studie geïdentifiseer was, was daar gepostuleer dat hierdie protease dieselfde is as die 29 kDa protease van *C. sonorensis*. Tans is daar baie min nukleotied reeks data van *C. imicola* beskikbaar en daarom is hierdie protease nie bekend nie. Die totale RNA van *C. imicola* was geïsoleer en daaruit was cDNA gesintetiseer. Die nukleotied reeks van *C. imicola* protease was ondersoek met behulp van die proses van polimerase ketting reaksie (PKR) en voorvoerders wat volgens die nukleotied reeks van die 29 kDa *C. sonorensis* protease ontwerp was. Aangesien die *C. sonorensis* protease 'n oop leesraam van 830 bp besit, was amplifisering van 'n sootgelyke grootte DNA fragment verwag. 'n Fragment met 'n oop leesraam van 830 bp was geamplifiseer. Met behulp van ‘n aanlyn soek enjin en BLAST is daar wel vasgestel dat die DNA nie vir enige bekende proteïene of proteases enkodeer nie.

Vanuit die literatuur wil dit wel voorkom asof daar 'n korrelasie is tussen die proteases in die muggie en soogdier gashere wat vatbaar is vir APS (Darpel *et al.,* 2011, Wilson *et al.,* 2009, Marchi *et al*., 1995). In hierdie studie kon ons vasstel dat daar ‘n proteïen in *C. imicola* is wat soortgelyk is aan die *C. sonorensis* 29 kDa proteïen. Hierdie proteïen is ook proteolities aktief. Verder was ‘n rekombinante 29 kDa *C. sonorensis* proteïen suksesvol uitgedruk en gedeeltelik gesuiwer. Hierdie nuwe kennis sal van waarde wees om in die toekoms beter insig te kry van die faktore wat bydra tot die oordrag van APSV, sowel as die wyses waarop proteolitiese veranderinge die infektiwiteit tussen verskillende gasheer spesies affekteer.

KEYWORDS

African horsesickness virus

Culicoides imicola

Culicoides sonorensis

Protease expression

Proteolytic activity

Recombinant protein expression

CHAPTER 1

Literature Review

1.1 Introduction

African horsesickness (AHS) is one of the most dreaded and deadly diseases of horses, which results in a mortality rate of over 90% in horses that have not been exposed to any African horsesickness virus (AHSV) serotype previously (Howell, 1963; Darpel *et al*., 2011). The disease is primarily endemic to sub-Saharan Africa and the Arabian Peninsula. However, several aspects of the epidemiology of AHSV indicate that it represents a significant risk to Europe and North America (Mellor *et al*., 1992). AHS is included in the 2013 notifiable disease list with an A list status (http://www.oie.int/animal-health-in-the-world/oie-listed-diseases-2013). AHS is a highly infectious and non-contagious arthropod-borne viral disease, which is recognised primarily as a pathogen of horses. Other equidae such as zebras and donkeys may also act as reservoirs for AHSV (Coetzer & Erasmus, 1994). This disease is caused by a member of the Orbivirus genus, within the family *Reoviridae* known as African horsesickness virus (Oellermann, 1970).

AHSV is a non-enveloped, double-layered icosahedrally symmetric virus with a linear double-stranded RNA (dsRNA) genome composed of ten segments (Oellermann, 1970). Each one of these ten genome segments code for a distinct viral protein, three non-structural (NS1, NS2 and NS3) and seven structural (VP1-VP7) (Bremer, 1976; Darpel *et al*., 2011). VP2 is the determinant for the virus serotype (Huismans *et al*., 1987). There are nine different AHSV serotypes (AHSV-1 to AHSV-9) (McIntosh, 1958). AHSV is also closely related to the prototype Orbivirus, namely Bluetongue virus (BTV), with many similar properties (Oellermann, 1970). Arboviruses are transmitted between mammalian hosts by arthropod vectors, which also become productively infected by these viruses. In certain cases the saliva from these arthropods can play an important role in the subsequent transmission of the virus to an uninfected host (Darpel *et al*., 2011;

Wilson *et al*., 2009). AHS viruses are transmitted between their mammalian hosts by certain haematophagous midge vectors (*Culicoides* spp.) (Erasmus, 1973).

Culicoides are biting midges within the family *Ceratopognidae* and are considered to be the most abundant of the haematophagous insects, as they can be found in most parts of the inhabited world (Mellor *et al*., 2000). The major vector that has been identified in the field of AHSV is *Culicoides* (*Avaritia*) *imicola* Kieffer (*C. imicola*) (du Toit, 1944), with studies suggesting *Culicoides bolitinos* (*C. bolitinos*)to be a possible second field vector (Meiswinkel *et al*., 1994; Mellor *et al.,* 2000). This is based on its geographical distribution, abundance in light traps and preference for larger mammals, especially horses (Mellor *et al.*, 1992). The North American BTV vector, *Culicoides sonorensis* (*C. sonorensis*) is also a highly efficient vector for AHSV under laboratory conditions (Boorman, 1974; Mellor *et al.*, 1975; Wellby *et al*., 1996). Furthermore, there have also been reports that not only vector proteases, but also the serum proteases of certain species are capable in cleaving the structural proteins of AHSV (Marchi *et al*., 1995; Darpel *et al.*, 2011). This cleavage could possibly contribute to an enhanced infectivity for the insect vector (Marchi *et al*., 1995).

Darpel *et al* (2011) identified a 29 kDa trypsin-like serine protease in the saliva of the BTV vector, *C. sonorensis,* which cleaves VP2 of the virus (Darpel *et al*., 2011). Previous studies performed on BTV and AHSV have also demonstrated proteolytic cleavage of the outer capsid protein, VP2, in the presence of trypsin or chymotrypsin (van Dijk & Huismans, 1982). This cleavage generated infectious sub-viral particles (ISVP) (van Dijk & Huismans, 1982; Marchi *et al*., 1995). Furthermore, there is also a possibility that AHSV particles circulating in susceptible species are ISVPs of which the outer capsid protein VP2 has been selectively cleaved by the action of trypsin-like serine proteases (Marchi *et al*., 1995). This enzymatic cleavage of virus particles resulted in enhanced viral infectivity of up to 1000-fold (Marchi *et al*., 1995). It is therefore important to determine what effect trypsin-like serine proteases will have on the viral proteins and infectivity of the AHSV particle.

1.2 Early history and epidemiology

Arnold Theiler discovered a filterable disease-causing agent in horses at Onderstepoort in South Africa in 1901 (Theiler, 1901). However, probably the first historical reference to AHS was in the Arabian document, 'Le Kitâb El-Akouá El-Káfiah Wa El Chafiâh' which reported an epizootic outbreak in Yemen in 1327 (Mouie, 1896). The AHSV is the cause of AHS in horses, donkeys, mules and sometimes even dogs (Alexander *et al.,* 1995; Coetzer & Erasmus, 1994; Mellor, 1994). Zebra species are important reservoir hosts, maintaining the virus in the field, but rarely displaying clinical signs of infection (Coetzer & Erasmus, 1994; Mellor, 1994). Depending on the virulence of the virus and the immune status of the infected horse, mortality rates of AHS can be up to 95% in horses (Coetzer & Erasmus, 1994; Wilson *et al.*, 2009).

Historically, despite small outbreaks in North Africa and the Arabian Peninsula, AHS was believed to be endemic to sub-Saharan Africa. However, from 1959 onwards, AHS has shown to expand beyond this core region with reported outbreaks in Pakistan, India, Spain and Portugal (Rafyi, 1961; Diaz & Panos, 1968; Mellor *et al.*, 1992). The first recorded outbreak of AHS in South Africa was in 1719 with the death of over 1700 horses in the Cape of Good Hope (Henning, 1956). Since then, different AHS outbreaks have been reported, highlighting the outbreak in South Africa that occurred from 1854 till 1855 in which over 70 000 horses died (Coetzer & Erasmus, 1994). However, the most significant recorded outbreak of AHS was in 1959 in the Middle East when over 300 000 horses died of the disease, resulting in virtual extinction of horses in that region. It was only with massive vaccination and vector control efforts that the disease was brought to a halt in 1961 (Howell, 1963). Small AHS outbreaks have persisted in geographical regions beyond that in which it usually occurs, suggesting a larger area suitable for sudden and rapid expansion (Mellor *et al*., 1992; Wilson *et al.*, 2009). However, the spread of the disease and possible outbreaks in countries such as America and Europe are currently prevented by prohibiting the importation of vaccinated horses or zebras from areas where the disease is endemic (Wilson *et al.*, 2009; MacLachlan & Guthrie, 2010).

Another virus from the same genus very prevalent and fatal in livestock was discovered by the chief veterinary officer of the Cape colony, Hutcheon, in the late 18th century (Paton, 1863). However, it was only in 1902 that this described disease was reported as bluetongue (BT) in the scientific literature for the first time (Hutcheon, 1902). BT, caused by the BTV, is a common livestock disease and is now found on almost all of the continents (Roy, 2001). BTV is the prototype Orbivirus and probably the most extensively studied virus within the Orbivirus genus. Furthermore, BTV shares a wide range of structural similarities with AHSV (Howell, 1962). Since AHS and BTV are some of the most lethal diseases of horses and cattle, both are included in the World Organisation for Animal Health's (OIE) list of notifiable diseases (OIE, 2013).

1.3 Pathogenesis of AHS

Depending on the form of AHSV infection in horses, clinical signs can vary greatly. Theiler described four different forms of the disease, namely the horsesickness fever form, the pulmonary (per-acute) form, the cardiac (sub-acute) form or the mixed (acute form) (Theiler, 1921).

The horsesickness fever form is the least severe form of AHS with the only clinical sign being a rise in body temperature to around 39°C - 40°C, which lasts about one to six days (Coetzer & Erasmus, 1994). A more severe form of AHS is the sub-acute or cardiac form. This form has a longer incubation period than the pulmonary form and clinical signs become visible around 7-12 days post infection. A characteristic feature of the cardiac form is the subcutaneous swelling of the head and neck, particularly the supraorbital fossae, palpebral conjuctiva and the intermandibular space (Figure 1.1). This form of the disease is also accompanied with a high fever ranging from 39°C - 41°C (Coetzer & Erasmus, 1994). The mortality rate is 50% - 70% and death usually occurs within four to eight days after the onset of the febrile reaction. Survivors recover typically in about seven days (Coetzer & Erasmus, 1994).

Figure 1.1 A horse before and during the cardiac form of AHS. Subcutaneous swelling of the supraorbital fossae, neck and intermandibular space is clearly visible (Seegers, 2010).

The pulmonary form of the disease is the most severe form with less than 5% of infected horses surviving. The pulmonary form most often occurs in horses with full susceptibility. The incubation period of the pulmonary form is relatively short and therefore clinical signs occur quickly following infection (Coetzer & Erasmus, 1994). The first clinical sign is a maximum fever of 41°C. The high fever is followed by severe dyspnoea and spasmodic coughing accompanied by large quantities of a frothy serofibrinous fluid discharged from the nostrils. Death usually occurs within a few hours after the disappearance of the severe dyspnoea (Coetzer & Erasmus, 1994). The most common occurring form of AHS is the mixed form. This form is characterised by a combination of high fever, respiratory distress and subcutaneous swelling, which ultimately results in the death of the infected horse (Coetzer & Erasmus, 1994).

There is currently no efficient form of treatment for AHS and this disease is kept in check by preventative methods such as vector control, quarantine in the case of an outbreak and vaccination. Vaccination is essential to protect horses against AHSV and to avoid new outbreaks. Currently, the only commercial vaccines for AHSV are live attenuated polyvalent and inactivated virus vaccines (Coetzer & Erasmus, 1994; Verwoerd & Erasmus, 1994). The attenuated live vaccine has proven to be highly effective but some drawbacks are still of concern. There is the

possibility that live attenuated vaccines can revert to virulence or recombine with wild type viruses (House, 1993b). The use of live attenuated vaccines also hampers the distinction between vaccinated and infected horses, which affects the international trade of horses (Martinez-Torrecuadrada *et al.*, 1996). Therefore, by better understanding the life cycle and molecular structure of the AHSV, effective and safe new generation vaccines that effectively prevent outbreaks could be developed.

1.4 AHSV Classification

The AHSV is an Orbivirus within the *Reoviridae* family of dsRNA viruses (Figure 1.2). Orbiviruses form one part of the six genera in the subfamily *Sedoreovirinae* and the word Orbivirus is derived from the latin word 'orbis', meaning ring. There are 22 serological groups of Orbiviruses and each of these groups consists of several serotypes (Mertens *et al.,* 1984). Even though Orbiviruses are clearly defined, they still form a large and diverse group within the *Reoviridae* family. Viruses within this family tend to show similar morphological and physiochemical properties (Gorman & Taylor, 1985; Urbano & Urbano, 1994). Well-known characteristics of the members of the *Reoviridae* family are their similar preference for vector hosts, mechanisms of transmission and geographical distribution (Urbano & Urbano, 1994).

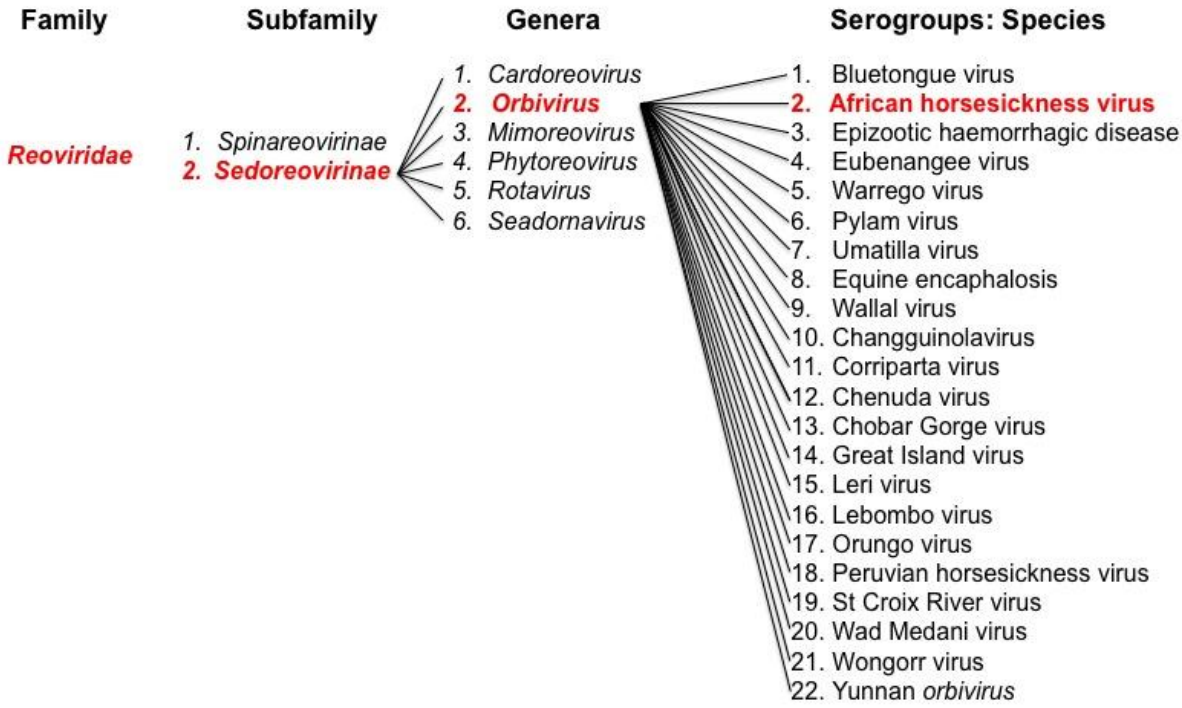

Figure 1.2 Classification of AHSV. AHS is an arthropod-borne viral disease of equids such as horses and is caused by the dsRNA-containing AHSV in the genus *Orbivirus*, of the subfamily *Sedoreovirinae* belonging to the family *Reoviridae*. This disease can be caused by any of the nine AHSV serotypes [International Committee on Taxonomy of Viruses (ICTV), 2013].

The main factor distinguishing Orbiviruses from other *Reoviridae* is the fact that they are capable of infecting both insects and vertebrates. The AHSV consists of nine serotypes (AHSV-1 to AHSV-9) (Howell, 1962; McIntosh, 1958) and has many significant similarities to BTV that has 24 serotypes (Howell, 1962; Gould & Hyatt, 1994).

1.5. Molecular biology of AHSV

1.5.1 AHSV Genome

The genome of Orbiviruses usually consists of ten dsRNA that are not lipid enveloped and are isometric in form. An inner core protects the dsRNA with a distinctive structure consisting of 32 ring-shaped capsomeres. These capsomeres are arranged in distinctive icosahedral symmetry (Nibert, 1998). Surrounding the

inner core is the outer shell that shows no distinctive capsomeric structure (Nibert, 1998). The genome of AHSV consists of ten segments of dsRNA, with each segment encoding at least one protein (Bremer, 1976). A feature that is very similar to most Orbiviruses is that each of the genome segments contains a long open reading frame that begins with AUG. This initiation codon is characteristically also protected from RNase degradation by the binding of ribosomes (Verwoerd, 1969). The RNA terminal sequences of all ten dsRNA segments of AHSV have conserved hexamers (Kiuchi *et al.,* 1983; Rao *et al.,* 1983). The 5’ end sequence of the positive mRNA strand is GUUAAA (Nibert *et al.*, 2001). The 3' conserved terminus is ACUUAC (Rao *et al.,* 1983; Kiuchi *et al.,* 1983; Mizukoshi *et al.,* 1993).

1.5.2 Viral morphology

The genome of BTV is probably the most extensively studied amongst the Orbiviruses. Although the genome of AHSV is distinct, it has not yet been studied as greatly as BTV. However, AHSV has many similar structural features with the genome of BTV (Figure 1.4) (Oellermann *et al.*, 1970; Bremer, 1976; Bremer *et al.,* 1990; Manole *et al*., 2012). Therefore, the morphology of AHSV can be explained based on the morphology of BTV. The core of BTV has an icosahedral nucleocapsid morphology with a core diameter of 69 nm (Figure 1.3) (Diprose *et al*., 2001).

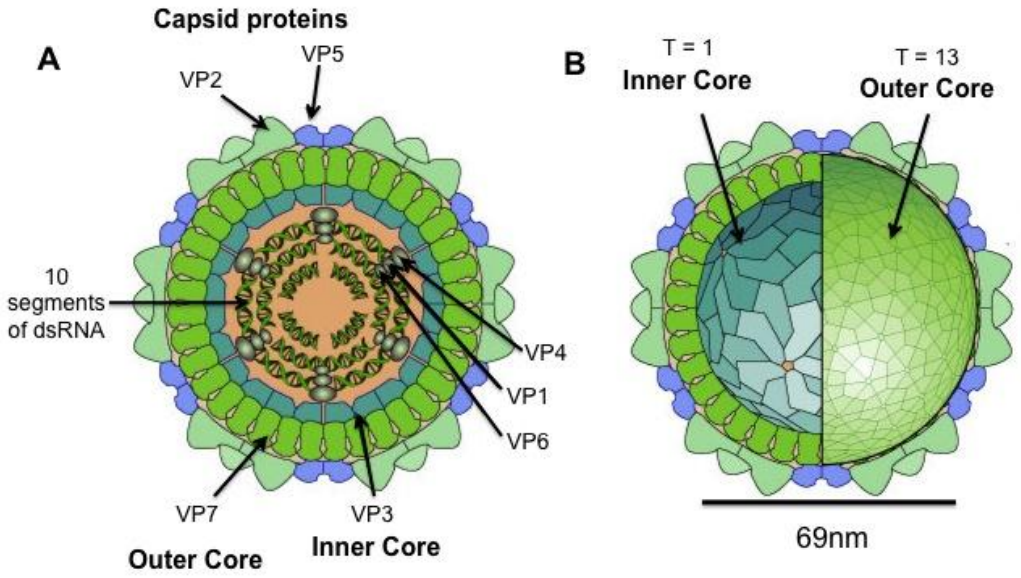

Figure 1.3 Diagram of the Orbivirus structure. Orbiviruses have non-enveloped, icosahedral, non-turreted virions with a triple-layered capsid structure of around 69 nm in core diameter. The inner capsid consists of viral proteins VP1, VP4 and VP6. The intermediate capsid consist of VP3 and VP7 and has a T=1 icosahedral symmetry. The outer capsid is comprised of VP5 and VP2 with a T=13 icosahedral symmetry (Swiss institute for bioinformatics, 2008; Mertens *et al*., 2005a).

Similar to BTV, the AHSV virion core consists of two protein shells, an inner and outer core layer (Owen & Mum, 1966), surrounding the ten dsRNA segments (Verwoerd *et al.,* 1972). The inner layer of the core has a T=1 symmetry and is made up of the minor proteins VP1, VP4 and VP6 (Prasad, 1992; Verwoerd *et al.,* 1972; Huismans & van Dijk, 1990). The outer layer of the core consists of the major proteins, 130 copies of VP3 and 780 copies of VP7 (Grubman & Lewis, 1992). This layer is formed by seven trimers, arranged in six five-membered rings. The five-membered rings are situated at the five-fold axis of the icosahedron (Roy *et al.,* 1997). VP7 consists of 260 trimers that covers VP3 (Diprose *et al*., 2001) and contains a tripod shaped upper and lower domain (Grimes *et al.,* 1995; Roy *et al.,* 1997). The disc shaped VP3 molecules has a T=13 symmetry and is responsible for forming the underlying smooth scaffold for the VP7 trimers (Prasad, 1992; Verwoerd *et al.,* 1972; Huismans & Van Dijk, 1990).

The outer capsid of the virion is comprised of two major proteins, namely VP5 and VP2 (Verwoerd *et al.,* 1972). The VP5 proteins, present as trimers, have a globular appearance and are underlying to the VP2 proteins (Roy, 2001). The sail-shaped VP2 proteins forms the outermost part of the virion capsid and almost completely

covers the VP7 trimers of the outer core layer (Mertens & Diprose, 2004). The outer capsid proteins, VP2 and VP5, form a continuous layer that, apart from the fivefold axis, completely covers the core layers of the virion (Verwoerd *et al.,* 1972).

Even though the structure of AHSV has not yet been studied as extensively as that of BTV, common features such as genome segment coding assignments, virion morphology and protein assembly are similar between AHSV and BTV (Fig 1.4 A-C) (Maree *et al.,* 1998; Bremer, 1976; Manole *et al*., 2012).

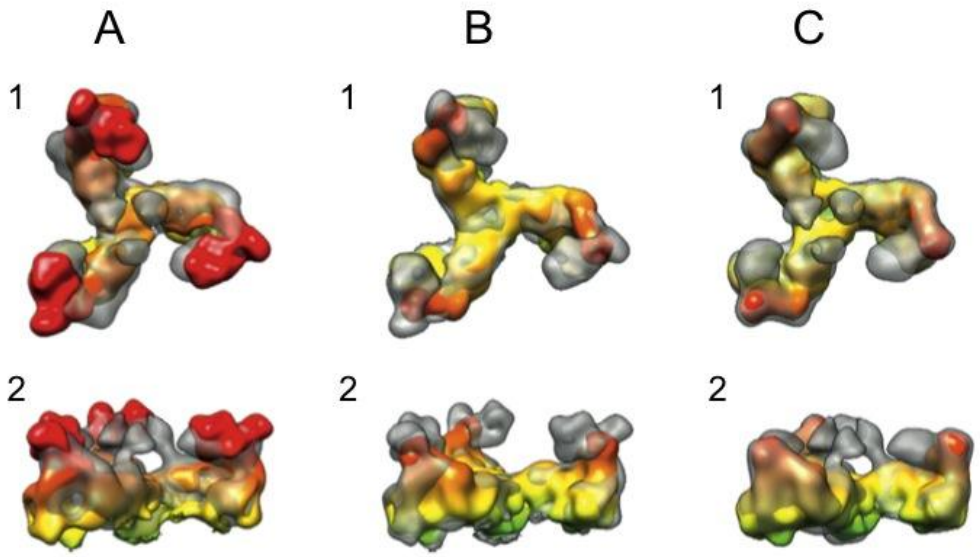

Figure 1.4 Structural comparisons between VP2 of AHSV4, BTV and AHSV7 with a truncated VP2. A) Composite figure of AHSV4 and BTV VP2 from the top (A1) and side (A2). The tip domains of BTV VP2 are indicated in red and the central domains of AHSV4 VP2 is indicated in transparent grey. B) Composite figure of AHSV7 with a truncated VP2 and BTV VP2 from the top (B1) and side (B2). The tip domains of BTV VP2 are absent at the truncated AHSV VP2 and are indicated in colour. The original structure of BTV VP2 is indicated in transparent grey. C) Composite figure of AHSV7 with a truncated VP2 and AHSV4 VP2 from the top (C1) and side (C2). The truncated AHSV7 VP2 is indicated in colour and the central domains of the original structure of AHSV4 in transparent grey (Manole *et al*., 2012).

1.5.3 AHSV Proteins

Each of the ten dsRNA genome segments of AHSV encodes for a specific protein. Seven genome segments encode for the structural proteins VP1 to VP7 and three encode for the non-structural proteins NS1, NS2, NS3 and NS3A (Grubman *et al.,* 1983; Mertens *et al.,* 1984; van Dijk & Huismans, 1988; Pedley *et al.,* 1988; Mertens *et al.*, 2005). The coding assignments of the genome segments as well as the functions of each of the proteins are summarised in Table 1.1.

Table 1.1. The AHSV genome segments, proteins they encode as well as the properties and functions of each protein (Mertens *et al.*, 2005; Maree *et al*., 1998b; Andrew *et al*., 1995)

Genome segment	Size (bp)	Encoded protein	Molecular weight (kDa)	Location in virion particle	Properties and Functions
1	3954	VP1	149	Inner core	RNA dependent RNA polymerase.
2	2926	VP2	120	Outer shell (Exposed)	Controls virus serotype; determine virulence; cleaved by proteases; cell attachment protein; haemagglutination; receptor binding.
3	2770	VP3	103	Sub-core layer (Scaffold)	Interacts with genomic RNA; forms scaffold for VP7 trimers; N-terminus is required for the binding and encapsidation of the transcription complex components; controls the size and organisation of the capsid structure.
4	1981	VP4	76	Inner core	Capping enzymes-guanylyltransferase, methyltransferases 1 & 2 and inorganic pyrophosphatase.
5	1638	VP5	59	Outer shell (Below surface)	Glycosylated; helps control virus serotype; causes membrane fusion indicating a role in virus penetration.
6	1769	NS1	64	Non-structural	Forms tubules of unknown function in the cytoplasm; minor immunogen for cytotoxic T-lymphocytes (CTL).
7	1156	VP7	38	Core surface layer	Core entry into cells; involved in infection of host and vector cells.
8	1124	NS2	41	Non-structural	Important viral inclusion body matrix protein; binds to single stranded RNA (ssRNA); phosphorylated and forms cytoplasmic inclusion bodies.
9	1046	VP6	35	Inner core	Binds ssRNA and dsRNA; helicase; ATPase
10	822	NS3/NS3A	25 / 24	Non-structural	Membrane protein; aids in virus release; glycoprotein; may be involved in determination of virulence.

1.5.4 Non-structural proteins

Besides the seven AHSV structural proteins used in the process of virus infection, three non-structural proteins also play a critical role. These non-structural proteins are better known as NS1, NS2 and NS3.

NS1 is encoded by genome segment six of the AHSV and has an estimated molecular weight of 64 kDa. A 95% NS1 amino acid identity has been observed amongst the different AHSV serotypes, indicating that the NS1 protein is highly conserved (Huismans & Els, 1979). Even though the function of NS1 is not yet fully understood (Mertens, 2004), there have been indications that NS1 assembles into tubules in infected cells (Huismans & Els, 1979; Maree & Huismans, 1997; Roy *et al.,* 1994b). These tubules formed by AHSV differ in appearance from the tubules formed by BTV NS1, since AHSV tubules have a 'cross hatch' internal structure with sharply defined edges and BTV a segmented ladder appearance (Huismans & Els, 1979; Urakawa & Roy, 1988; Maree & Huismans, 1997).

The non-structural protein, NS2, is encoded by genome segment eight and has a molecular weight of 41 kDa. Cellular kinases of NS2 play an important role in the ability of the virus to bind to ssRNA (Huismans *et al.,* 1987b; Theron & Nel, 1997), and may therefore be important in the formation of virus inclusion bodies (Roy *et al.,* 1994b) and controlling the AHSV replication in cells (Theron *et al.,* 1994). There have also been indications that NS2 and NS3 are involved in the assembly and release of the virus (Roy *et al.,* 1994b; Mertens, 2004).

There are two non-structural proteins that aid in the release of the virus, namely NS3 and NS3A (Van Staden & Huismans, 1991; Mertens, 2004). NS3 and NS3A are membrane proteins encoded by genome segment ten and have molecular weights of approximately 25 kDa and 24 kDa respectively. NS3 is responsible for the binding of shelled viruses facilitating budding from the endoplasmic reticulum for virus maturation (Mertens, 2004; Van Staden *et al.,* 1995) and possible release of the virus from the cell (Roy *et al.,* 1994b). Studies have also shown that NS3 and NS3A is a minor immunogen for CTL (Andrew *et al*., 1995; Mertens, 2004).

1.5.5 Minor and major core proteins

Each of the ten dsRNA segments of the AHSV is particularly organised within the core and arranged around a transcription complex situated at the fivefold axis. The core of AHSV is made up of a maximum of 12 transcription complexes (Mertens & Diprose, 2004) consisting of core proteins with assigned enzymatic activity. The TCs consist of VP1 polymerase (Roy *et al.,* 1988), the VP4 capping enzyme (Ramadevi *et al.,* 1998) and the VP6 helicase (Stauber *et al.,* 1997).

The largest genome segment of AHSV encodes VP1, which has an estimated molecular weight of around 149 kDa (Vreede & Huismans, 1998). Studies performed on the function of VP1 suggest RNA-dependent RNA polymerase activity, equivalent to the VP1 protein of rotaviruses (Vreede & Huismans, 1998; Both *et al.,* 1994). Genome segment four of the AHSV encodes VP4, a protein with a molecular weight of around 76 kDa. VP4 contains the capping enzymes guanylyl transferase, methyltransferases 1 & 2, NTPase and inorganic pyrophosphatases (Roy *et aI.,* 1994b). Dimers are also formed by VP4 of the structurally similar Orbivirus, BTV, through a leucine zipper, contributing to the assembly of the core particle (Ramadevi *et aI.,* 1998). Genome segment nine of AHSV encodes the structural protein VP6 and has a predicted size of around 35 kDa. The major function of VP6 is the binding of ssRNA and dsRNA, which in return, stimulate ATPase function and plays a major role in the assembly of the virus core (Stauber *et al.,* 1997*).*

The innermost major core protein of AHSV, VP3, is encoded by genome segment three and has a molecular weight of around 103 kDa. VP3 forms interactions with the genomic RNA and contains group specific antigenic determinants (Roy *et al.,* 1994b, Roy & Sutton, 1998). The binding and encapsidation of the transcription complexes components relies on the N-terminus of VP3. The VP3 protein has chemically featureless grooves on the inside that forms tracks for the RNA and also plays a fundamental role in assembling the core (Kar *et al.,* 2004). The VP3 dimers can be seen as building blocks for the icosahedral structure of the sub-core as it forms scaffolding for VP5 and VP2 proteins to bind to (Roy, 1992; Manole *et al*., 2012).

VP7 forms the outermost major core protein of the AHSV and is encoded by genome segment seven. The VP7 protein has a predicted molecular weight of around 38 kDa, is highly conserved amongst Orbiviruses and is also rich in hydrophobic amino acids. VP7 also helps with the stabalisation of the innermost major core (Mellor & Mertens, 2008). The protein structure and its function have been extensively studied (Grimes *et al.,* 1998; Manole *et al*., 2012), demonstrating the ability to induce an immune response in mice during an AHSV serotype challenge (Wade-Evans *et al.,* 1997). This study confirmed that VP7 is a group specific antigenic determinant (Huismans *et al*., 1987; Roy *et al.,* 1994b), although further research is necessary to determine the exact functions of AHSV VP7.

1.5.6 Major capsid proteins

The major capsid proteins of AHSV, VP5 and VP2 are the most variable of the viral proteins. The VP5 protein is encoded by genome segment five and has a molecular weight of around 59 kDa. Although the VP5 protein is not nearly as exposed to the outside as VP2 (Roy *et al.,* 1994b; Manole *et al*., 2012), VP5 possibly interacts with VP2 and thereby enhances the immune response against AHSV (Martinez-Torrecuadrada *et al.,* 1996; Mertens *et al.,* 1989; Darpel *et al*., 2012). However, VP5 does not seem to be essential for eliciting a protective immune response (Huismans *et al.,* 1987a, Martinez-Torrecuadrada *et al.,* 1996). The VP5 protein is highly conserved amongst the different AHSV serotypes with an amino acid identity of 96% between African horsesickness serotype 9 (AHSV9) and AHSV4 (du Plessis & Nel 1997; Kaname *et al.*, 2013). O'Hara *et al* (1998) established the importance of the VP5 protein to the AHSV virion through the suggestion that VP5 is linked to the determination of virulence (O'Hara *et al.,* 1998).

VP2 is the most variable protein of the virion and shares only 19-24% of identical VP2 amino acids with other Orbivirus species (Iwata *et al.,* 1992; Kaname *et al*., 2013). VP2 is encoded by genome segment two of the AHSV and is around 120 kDa in size (Roy *et al.,* 1994b). The C-terminus of VP2 seems to be the most conserved and is possibly the region that forms interactions with VP5 and VP7 (Roy *et al.,* 1994b; Vreede & Huismans, 1994; Kaname *et al*., 2013). The VP2

protein is a target for neutralising monoclonal antibodies generated during infection of the mammalian host (Burrage & Laegreid, 1994), and can therefore be seen as the major serotype determinant of an individual strain (Huismanns, 1987; Kaname *et al*., 2013). It was also demonstrated that VP2 induces protection against AHSV serotype specific infection during immunological experiments (Martinez-Torrecuadrada *et al.,* 1994; Roy *et al.,* 1996; Darpel *et al*., 2011). Studies have confirmed that VP2 is very sensitive to equine serum proteases and that proteolytic cleavage of the BTV VP2 in the saliva of the vector results in an increased infectivity of the virus (Burroughs, 1994; Marchi *et al.,* 1995; Darpel *et al*., 2011). This may suggest that VP2 plays an important role in determining virulence (O'Hara *et al.,* 1998; Darpel *et al*., 2011).

It is not yet clearly understood how the proteolytic cleavage of VP2 affects the infectivity of the virus. However, is known that VP2 plays an important role in the initial steps of infection. Therefore, by gaining better insight into the factors that influence VP2 in such a manner that it increases infectivity is important. This could aid in future development of possible vaccines (Castillo-olivares, 2011; Kaname *et al*., 2013).

1.6 AHSV transmission and replication

1.6.1 AHSV transmission

Pitchford and Theiler proposed in 1903 that biting insects from the genus *Culicoides* transmit AHSV (Pitchford & Theiler, 1903). However, the first definitive biological transmission of AHSV through bites of the haematophagous athropods, *Culicoides*, was only demonstrated four decades later (du Toit, 1944). Apart from the ingestion of virus-contaminated meat by dogs (Van Rensburg *et al.*, 1981; Hess, 1988), AHSV transmission is virtually exclusively controlled by the arthropod vector. To date, the only two domestic species known to be extremely susceptible to AHSV infection are horses and dogs and these animals suffer high mortality rates during epidemics of the disease (Dardiri & Salama, 1988; Van Rensburg *et*

al., 1981). The basic biological process of AHS transmission is depicted in Figure 1.5.

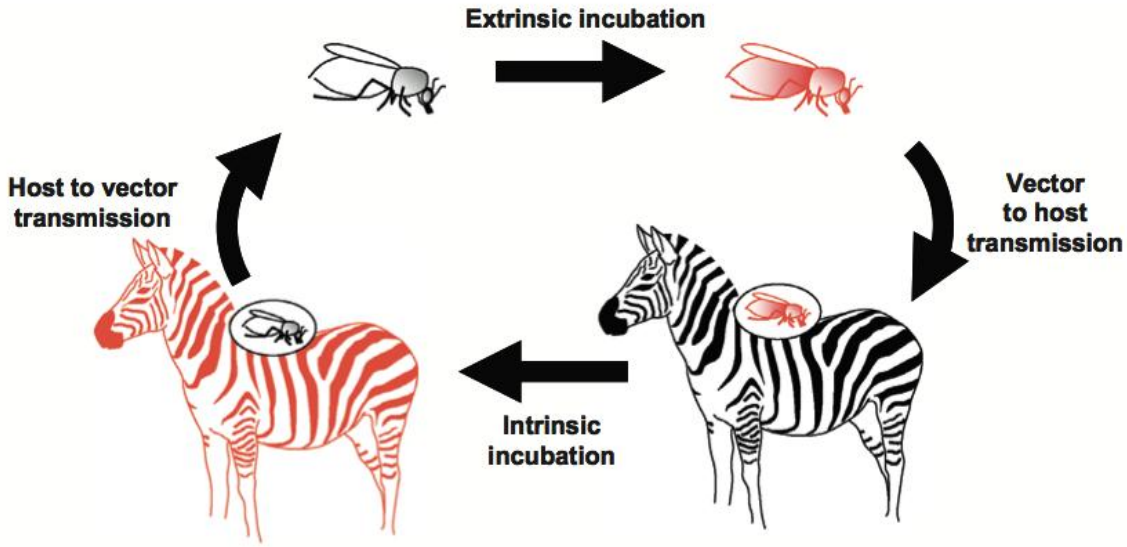

Figure 1.5 The AHSV transmission cycle. The *Culicoides* vector feeds on an AHSV infected vertebrate. The ingested viruses are transported to the gut of the midge where the virus replicates. Mature virions migrate to the salivary glands and are inoculated into a susceptible vertebrate host through the process of biting and feeding. The inoculated virus will replicate within the infected host and mature, causing the characteristic pathological features of AHS. The cycle will subsequently repeat itself by midges feeding on the infected host (Wilson *et al*., 2009).

An AHSV infected vertebrate host such as a horse or zebra will have virus circulating in its skin tissue or peripheral blood vessels (Wilson *et al.,* 2009). This makes the virus accessible to the blood feeding arthropods. *Culicoides* species are pool feeders, meaning they use their rasping proboscis in the presence of saliva to create a small wound in the skin. This is followed by the uptake of the material influx, which includes inoculated saliva, lymph and virus containing blood (Hocking, 1971). From hereon, the virus will be transported to the mesenteronal lumen of the gut where it will penetrate and infect the cells of the gut wall. The virus will then enter the haemocoel and slowly migrate through the internal environment of the midge targeting the secondary organs of infection such as the salivary glands (Mellor *et al*., 2000; Wilson *et al.,* 2009).

The skin of the host acts as a mechanical and immunological barrier (Wikel, 1996). However, studies have established that the saliva of arthropods contains

components capable of inhibiting the immunological response of the host skin (Gillespie *et al*., 2000) and promoting viral replication (Darpel *et al.*, 2011). It is also suggested that components in the saliva of the *Culicoides* species are capable of modulating the local blood flow at the feeding site of the host (Titus & Ribeiro, 1990). Upon feeding, the virus is transmitted through the virus-infected saliva back to the vertebrate host. Here, the virus will enter the bloodstream and spread to the different susceptible organs, ultimately resulting in the characteristic pathological features of AHS. The transmission cycle of AHSV is dependent on two forms of incubation namely, intrinsic and extrinsic incubation. Intrinsic incubation is the time that it takes between ingestion of the virus by the midge and the midge being able to transmit the virus to another vertebrate host. Extrinsic incubation is based on external factors affecting the vectors' ability to transmit the virus such as environmental temperature (Wilson *et al.,* 2009).

1.6.2 Infective replication cycle of BTV

Even though AHSV and BTV are distinct at genetic level and in terms of the proteins they encode, they share similarities in morphology and coding strategies. Therefore, the infective replication cycle of AHSV is speculated to be similar to the replication cycle of BTV (Mertens & Diprose, 2004; Matsuo & Roy, 2009; Kaname *et al*., 2013). Figure 1.6 depicts the infective replication cycle of BTV.

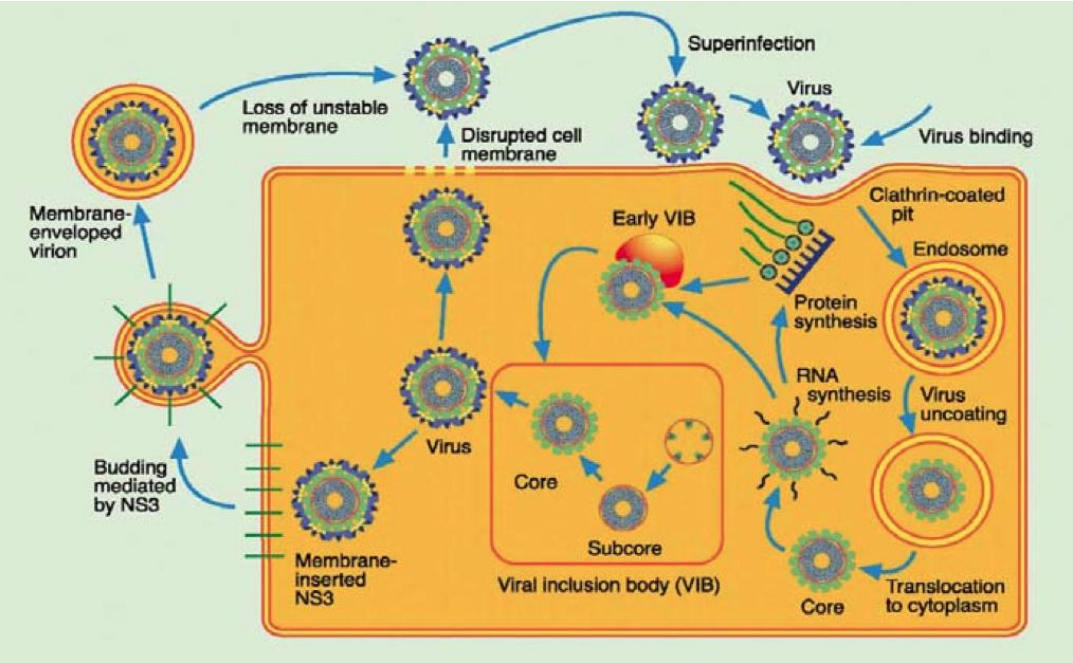

Figure 1.6 Schematic diagram representing the infective replication cycle of bluetongue virus (Mertens & Diprose, 2004).

Following transmission, the larger capsid protein, VP2, binds to the outer surface of the susceptible hosts' cells and the infecting viral particles are invaginated through the clathrin-mediated process of endocytosis (Hyatt et al 1993; Mertens & Diprose, 2004). This results in the formation of an endosome vesicle containing the virus. Detachment of the endosome from the cell surface releases the endosome into the cell plasma. The endosome has a low pH, which aids in uncoating of the outer capsid proteins VP2 and VP5 from the virus core and thereby possibly increasing the infectivity of the viral particle (Hutchinson, 1999; Roy, 2001; Mertens *et al.*, 2004; Forzan *et al*., 2007). The removal of the outer capsid proteins within one hour of infection is very important for the activation of the core-associated RNA-dependent RNA polymerase (Mertens *et al.,* 2004; Verwoerd *et al.,* 1972; Huismans *et al.,* 1987b). The release of VP2 and VP5 results in VP7 being exposed, and plays an essential part in the translocation of the core particle from the endosome into the cytoplasm of the host cell (Mertens & Diprose, 2004; Forzan *et al.,* 2007; Mertens *et al*., 2004). The transcriptase functions of the core particle are activated during the uncoating of the outer capsid proteins. The viral polymerase, VP1, synthesises a full-length positive stranded mRNA transcript from each of the dsRNA segments that is capped and methylated while they are translocated from the core through the pores at the fivefold axis into

the cytoplasm (Huismans and Verwoerd, 1973; Roy, 2001; Mertens & Diprose, 2004). These core particles have two major functions namely, it provides the templates for the synthesis of the new dsRNA genome and provides the mRNA for the synthesis of viral proteins (Matsuo & Roy, 2009). During the process of mRNA synthesis, the dsRNA is never exposed to the cytoplasm. This prevents the activation of the host defense mechanisms in response to the dsRNA (Mertens *et al*., 2004). In the cytoplasm, the newly synthesised mRNA transcripts are used as templates for translation. Here, NS2 plays a major role in selecting the distinct mRNA segments for the assembly of the newly formed viral proteins into sub-viral core particles (Roy, 2001; Mertens & Diprose, 2004). Inclusion bodies primarily consisting of NS2, surrounds the cores of these newly formed sub-viral core particles (Huismans *et al.,* 1987a). These inclusion bodies serve as the environment in which the progeny viruses are subsequently formed (Lecatsas, 1968; Roy, 2001; Mertens & Diprose, 2004). The capsid proteins VP2 and VP5 are located at the periphery of the inclusion bodies and are assembled onto the sub-viral core particle to form a mature viral particle. The mature viral particles leave the cell in one of two ways. Either by the action of the integral membrane protein, NS3, which mediates the release of the newly synthesised virus particle by means of budding through the cell membrane or following cell death associated with the local disruption of the cell plasma membrane (Van Staden & Huismans, 1991; Roy, 2001).

1.7 Vector species of AHSV

Culicoides are biting midges within the family *Ceratopognidae* and are considered to be the most abundant of the haematophagous insects, as they can be found in most parts of the inhabited world (Mellor *et al*., 2000) (Figure 1.7). *Culicoides* are the world's smallest heamatophagous midges measuring only between 1-3 mm in size (Mellor *et al*., 2000). These midges are known to transmit a great variety of pathogens between domestic and wild animals and in some cases even humans. But, it was the ability of *Culicoides* to act as a vector for Arboviruses of domestic livestock that has ensured them a position in the spotlight (Mellor *et al*., 2000). The

species of *Culicoides* are abundant in most regions of the world except in New Zealand, the Hawaiian Islands and extreme Polar regions (Mellor *et al*., 2000).

Figure 1.7 The vector of AHSV, *Culicoides* (*Avaritia*) *imicola* Kiefer (Mathieu, B. 2010).

The first time *Culicoides* was described in the literature was in 1713, when Derham wrote about their biting habits and life history (Derham, 1713). Since then, over a 100 species of blood feeding *Culicoides* have been identified through collections worldwide. Around 22 species can be found near livestock. A collection of pictures containing the medically and veterinary important species of *Culicoides* was compiled by Boorman (1993). These pictures focused on the specific wing patterns of *Culicoides,* which are used to differentiate between the different species (Boorman, 1993). The lifespan of *Culicoides* is considered to be relatively short, as individuals survive at most only 20 days (Blanton & Wirth, 1979). Since the adult *Culicoides* are nocturnal and crepuscular, activity peaks around sunset and sunrise with a decline in activity during the night and hardly any activity during daytime (Venter *et al*., 2009; Mellor *et al*., 2000). The differences between the biology of the *Culicoides* species is vast, with not all of them feeding on mammals or even being susceptible virus infections (Mellor *et al*., 2000). Based on the feeding preferences of these midges, *Culicoides* have the ability to potentially transmit diseases such as AHS and BTV (Mellor *et al*., 2000). Throughout the world, a variety of viruses have been isolated from the *Culicoides* species with 19 types belonging to the family *Reoviridae* (Meiswinkel *et al*., 1994).

The only confirmed field vector of AHSV is the biting midge *C. imicola*, although *C. bolitinos* is also a suspected vector (Meiswinkel *et al.,* 1994; *Bouayoune et al.*, 1998; Mellor *et al.,* 2000). Numbers of up to 1,5 million *Culicoides* can be collected in light traps near livestock in a single night (Van Ark & Meiswinkel, 1992). These collections indicated that *C. imicola* is the most abundant livestock-associated *Culicoides* species in South Africa (Mellor *et al*., 2000). *C. imicola* is considered to be the major vector for AHSV in South Africa (Nevill *et al*., 1992), however, various isolations of the virus have also been made in Zimbabwe (Blackburn *et al.,* 1985), Kenya (Walker & Davies, 1971) and even as far as the Sudan (Mellor *et al*., 1984). It is believed that *C. imicola* has a year-round presence during the adult phase, which increases the possibility of these regions to be enzootic zones for AHS (Rawlings *et al*., 1997; Mellor *et al*., 2000). *C. imicola* prefer summer rainfall regions in the north of the country with a warm frost-free climate, which makes these regions excellent AHS endemic areas (Mellor *et al*., 2000). However, it has been established that *C. imicola* has an aversion for very warm sandy regions such as the Karoo as well as cooler high lying areas of the country (Mellor *et al*., 2000). Studies suggest that the possible AHSV vector, *C. bolitinos*, favours these cooler regions (Mellor *et al*., 2000), which could possibly result in a larger area becoming endemic for AHSV.

Upon feeding on the blood of the vertebrate host, the viruses are either transmitted from the midge to the blood stream of the vertebrate host or the midge ingests viraemic blood from the vertebrate host (Wilson, 2009; Mellor *et al*., 2000). Only the female *Culicoides* prefer to feed on larger animals, as they need to obtain a good protein source for egg production (Wilson, 2009). Higher ambient temperatures are needed for the production of eggs, which can be as quickly as every 3-5 days. Lower ambient temperatures will result in a decrease in time between blood meals as well as replication of the virus. Since the *Culicoides* are incapable of regulating their body temperature, their life cycle and egg production is directly dependent on environmental temperature, but even at lower temperatures, the midges are still capable of transmitting the virus (Mellor *et al*., 2000).

As of yet, not much is known about the preferred breeding locations of *C. imicola*, but is believed that they primarily breed in unvegetated areas covered in short grass, which is located close to water with organically enriched and damp soil (Nevill, 1972; Mellor *et al*., 2000). Equids and ruminants situated near these areas will therefore enable the midges to opportunistically feed on them and subsequently contribute to the transmission the virus (Nevill, 1972).

A second possible vector species for AHSV, *C. bolitinos*, has recently been shown to be more and more susceptible to AHSV infections, with successful infections in the laboratory and AHSV isolations from field-caught specimens during AHS outbreaks (Wilson *et al.,* 2009; Mellor *et al*., 2000). Similar to *C. imicola*, *C. bolitinos* also feed on larger livestock such as horses, but is generally up to ten times less abundant than *C. imicola* and is geographically present in cooler regions of South Africa (Van Ark & Meiswinkel, 1992; Mellor *et al*., 2000). Laboratory studies confirmed *C. bolitinos* to be a competent vector for AHSV as it was susceptible to oral infection. This finding was the same as with *C. imicola*, which supports the possibly of *C. bolitinos* to be a second field vector of AHS (Mellor *et al*., 2000). Although *C. imicola* is considered to be the primary vector for the transmission of AHSV, it has been reported that the North American BTV vector, *C. sonorensis*, is a highly efficient vector for AHSV under laboratory conditions (Mellor, 1975; Boorman *et al*., 1975; Wellby *et al*., 1996). Studies based on the ability of AHSV to infect and replicate within *Culicoides* species have shown that at least 11 of the 22 livestock-associated *Culicoides* species are capable of being a successful host for AHSV (Mellor *et al*., 2000).

Based on the abundance and prevalence of the midges, it is nearly impossible to effectively control these insect vectors. Therefore, research should rather concentrate on understanding the mechanisms implemented by these vectors that aid the transmission of AHSV. This might contribute to future development of methods that could help with the protection of susceptible hosts as well as predicting when and where future disease outbreaks may occur.

1.8 Proteolytic cleavage of VP2

AHSV VP2 is an important determinant of the vector immune response, since it elicits neutralising antibodies in vaccinated and infected animals. VP2 is very sensitive to equine serum proteases (Burroughs *et al.,* 1994), but the salivary proteins of haematophagous arthropods have also been demonstrated to play a critical role in the initial steps of infection and mediate pathogen transmission (Ribeiro & Francischetti, 2003; Langer *et al*., 2007).

Mertens *et al* (1987) demonstrated that the outer capsid protein of AHSV and BTV, VP2, could be directly modified through proteolytic cleavage with trypsin or chymotrypsin (Mertens *et al.,* 1987; Burroughs *et al.,* 1994). The proteolytic enzymes within the saliva of the arthropod vectors can cleave the Orbivirus VP2 (Darpel *et al*., 2011), which is responsible for haemagglutination (Van der Walt, 1980). This will result in an increase in viral infectivity in midges (Marchi *et al.*, 1995). A study performed by Darpel *et al* (2011) demonstrated that proteases are present in the saliva of the BTV vector, *C. sonorensis*, which yielded nearly identical VP2 cleavage patterns when compared to trypsin digestion (Darpel *et al*., 2011; Van Dijk & Huismans, 1982). Therefore, it is highly likely that endoproteolytic cleavage of VP2 occurs in the saliva of the *Culicoides* host. A common post-translational modification of viral membrane proteins is endoproteolytic cleavage, which usually occurs at the arginine residues. This can be seen as a crucial factor in the steps for determining the pathogenicity of Orbiviruses (Klenk & Garten, 1994). Since VP2 is responsible for haemagglutination and virus entry (Van der Walt, 1980), the cleavage of VP2 by the adult female *Culicoides* salivary proteases could possibly lead to an enhanced infectivity of the virus particle during the initial stages of infection (Darpel *et al*., 2011).

The VP2 amino acids ranging from position 340 to 360 of AHSV have previously been speculated to determine virulence and tissue tropism (Potgieter *et al*., 2009; Manole *et al*., 2012). Manole *et al* (2012) reported the discovery of an AHSV strain with a naturally occurring truncated VP2 that had a deletion of 225 amino acids

ranging from residues 279 to 503 (Figure 1.8) (Manole *et al*., 2012).

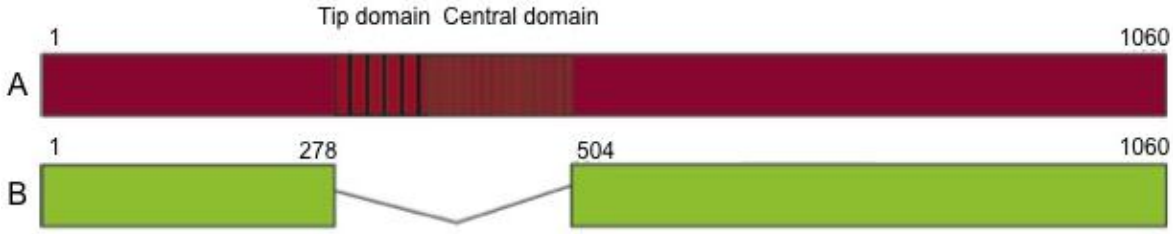

Figure 1.8 Comparison between deletions in AHSV VP2 and AHSV4 tVP2. Figure A is a schematic of the full sequence of AHSV4 VP2. Figure B shows the major deletions from residues 279 to 503 in AHSV7 tVP2. The figures were composed based on multiple sequence alignments (Manole *et al*., 2012)

This deletion is present in a region that is known to determine virulence and contains immunogenic epitopes (Manole *et al*., 2012; Martinez-Torrecuadrada & Casal, 1995). Immediately following inoculation, proteolytic enzymes will modify the virus particles generated by replication within the salivary glands of the *Culicoides* vector (Darpel *et al*., 2011). The selective cleavage of the outer capsid protein VP2 through the action of trypsin-like proteases changes the structure of the virus, which results in the formation of infectious sub-viral particles. These infectious sub-viral particles have been demonstrated to play a highly significant role during the initial stages of infection by increasing the specific infectivity of the virus particle in midges (Dimmock, 1982; Marchi *et al*., 1995; Darpel *et al*., 2011). Furthermore, when comparing the proteolytically cleaved VP2 of AHSV and BTV with intact virus particles, an increase of up to a 1000-fold in infectivity for the viral particles can be seen in the adult *Culicoides* (Mertens *et al*., 1987; Darpel *et al*., 2011). Manole *et al* (2012) also demonstrated that the AHSV7 strain with the naturally truncated VP2 was more infective than the AHSV4 with a full length VP2. The AHSV with the truncated VP2 showed full cytopathic effect in tissue culture much faster and also outgrew the AHSV4 with the unaltered VP2 (Manole *et al*., 2012). Furthermore, Darpel *et al* (2011) demonstrated that the specific infectivity of BTV particles for KC cells increased by up to ten-fold when treating the virus particles with salivary proteases of *C. sonorensis* (Darpel *et al*., 2011) suggesting a similar cleavage mechanism of VP2 that resulted in the truncation of the AHSV VP2 observed by Manole *et al* (2012).

Therefore, proteolytic cleavage of VP2 may be important in determining selective growth advantage and the infection rates in insect vectors as well as mammalian hosts. Since the central domain of VP2 is located on the top of the putative sialic acid binding site (Zhang *et al*., 2010), it is assumed that the central domain of VP2 functions as a good potential target for the proteases in the saliva of the *Culicoides* vector and the serum in the susceptible vertebrate hosts (Burroughs, 1994; Marchi, 1995). This can possibly be based on the absence of this central domain region in AHSV with the truncated AHSV7 VP2, which have shown an increased infectivity and growth rate compared to AHSV with a non-truncated VP2 (Manole *et al*., 2012). Intermediate breakdown products are initially formed by the proteases of vector species, but it is believed that this initial cleavage does not in itself alter the infectivity of the virus. Therefore, additional cleavage or processing of VP2 by the serum proteases of the vertebrate host might be needed in order to fully achieve increased specific infectivity (Darpel *et al*., 2011). However, up untill now there has been no direct evidence that could link VP2 cleavage of AHSV or BTV to an increased infectivity in mammalian hosts (Marchi *et al*., 1995; Darpel *et al*., 2011). The efficiency of VP2 cleavage by these proteases is believed to be dependent on a variety of factors such as the protein composition of the vector saliva, the ability of vector proteases to digest VP2, the susceptibility of the host cells to the cleaved VP2, pH and temperature (Wilson *et al*., 2009; Darpel *et al*., 2011). These factors play a major role in the initial stages of virus-cell interactions and infection of the insect midgut (Darpel *et al*., 2011). Different Orbiviruses have different VP2 cleavage sites for the vector proteases. However, the ability of different *Culicoides* proteases to efficiently cleave VP2, as well as their cleavage efficiency between the different *Culicoides* species is to date still unclear.

The exact mechanisms behind VP2 cleavage that possibly lead to an increased virulence is not yet understood well, but it appears that the virus particle may have adapted to make use of these changes to have an enhanced ability to infect *Culicoides* cells by means of saliva activated transmission (Wilson *et al.*, 2009; Laegreid *et al.,* 1993).

1.9 Problem formulation and aims of this study

Even though AHSV is currently confined to sub-Saharan Africa, it has shown the ability to expand beyond the current core region on several occasions. This demonstrates that certain aspects of AHSV epidemiology do represent a significant risk to countries outside of this region, affecting the worldwide equine industry. Darpel *et al* (2011) demonstrated that saliva proteases of the *Culicoides* vector play an important role in the transmission of the Orbivirus prototype, BTV (Darpel *et al*., 2011). The trypsin-like serine salivary protease in *C. sonorensis* is able to cleave VP2 of BTV that results in an increased capacity of the virus to infect cells (Darpel *et al*., 2011). Manole *et al* (2012) identified a naturally occurring AHSV strain with a truncated VP2 that showed similar infection abilities when compared to the protease treated BTV VP2 of Darpel *et al* (2011) (Manole *et al*., 2012; Darpel *et al*., 2011). Therefore, it seems as if a correlation can be made between protease digested VP2 and an increase in infectivity. The importance of modifications of the virus particle with saliva proteases that ultimately aids in virus transmission is not yet well understood. It is important to better understand the factors that contribute to Orbivirus transmission and how proteolytic viral modifications will affect the capacity of the virus to infect different host species and cause clinical signs.

The overall objective for this study was to start to investigate the *Culicoides* salivary proteases associated with altering VP2 of AHSV locally in such a way that it increases the infectivity of the virus. In order to achieve this goal, the different objectives for this study were formulated as follows:

1. To determine if proteases are present in the total homogenate of the vector for AHSV and BTV, *C. imicola*, and if these proteases were characteristically similar to the salivary proteases of *C. sonorensis*
2. To express a recombinant salivary trypsin-like serine protease identified by Darpel *et al* (2011) in *C. sonorensis*, to purify the recombinant protease using metal chelate affinity chromatography and investigate the proteolytic character of the purified protease with the use of different protease inhibitors and substrates

3. To incubate purified AHSV with the purified recombinant protease to identify possible VP2 cleavage sites
4. To identify a 29 kDa protease in *C. imicola* similar to the 29 kDa protease identified in *C. sonorensis* by performing reverse transcriptase polymerase chain reaction (RT-PCR) on synthesised *C. imicola* complimentary DNA (cDNA) and using primers based on the nucleic acid sequence encoding the *C. sonorensis* trypsin-like serine protease

CHAPTER 2

Detection of proteases in the total protein extract of *Culicoides imicola*

2.1 Introduction

Culicoides imicola is known to be one of the major vectors contributing to the transmission of AHSV (Nevill *et al*., 1992). Even though *C. sonorensis* primarily transmits BTV, it has been reported to be able to transmit AHSV in the laboratory (Boorman, 1974). Furthermore, to date, different *Culicoides* species have been associated with the transmission of different Orbiviruses. This suggests that different mechanisms of viral transmission are most likely implemented by the different vector species. Supporting this statement, studies by Mertens *et al* (1996) and Darpel *et al* (2011) have also reported that a protein in the saliva of *C. sonorensis* has trypsin-like characteristics and is responsible for cleaving VP2 of BTV resulting in an increased infectivity of the virus. Our long-term goal is to gain more knowledge about the role of proteases in the replication cycle of AHSV and virus-vector host transmission. The structure of AHS is very similar to BTV (Manole *et al*., 2012). Based on this, it is important to identify and study the proteases in *C. imicola* possibly responsible for digesting the outer capsid protein VP2 of AHSV. The primary aims of this part of the study were to extract the total proteins present in *C. imicola*, to determine if there are active proteases amongst these proteins and to establish if these are trypsin-like proteins.

To determine if *C. imicola* has a protease similar in size to the 29 kDa protease identified by Darpel *et al* (2011) without the tedious process of collecting vector saliva. Instead, the experimental approach was to homogenise the *Culicoides* insects and separate the total proteins present in the homogenate by sodium dodecyl sulfate polyacryamide gel electrophoresis (SDS-PAGE). To determine if these proteins have proteolytic activity, the idea was to by subject these proteins to a proteolytic activity assay. Further characterisation was to be done by means

of incubation in the presence of different class-specific protease inhibitors to determine to what class of protease the activity belongs to.

2.2 Materials and Methods

All of the materials and reagents used in this study, along with the suppliers and product numbers are supplied in Appendix I. To simplify the flow of the text, the supplier and catalogue number will not be mentioned in the text.

2.2.1 Collection and sub-sampling of *C. imicola*

The biting midges of the species *C. imicola*, used in this part of this study was generously collected and sorted by Dr. G.J. Venter and his team at the Agricultural Research Council- Onderstepoort Veterinary Institute (ARC-OVI), Pretoria, South Africa. The trap used for collection (Figure 2.1) is a modified version of an insect trap that was commercially available in Europe in the early 1970s. The trap is now manufactured in South Africa and primarily used in *Culicoides* surveys. The 220 V trap functions on the basis of downdraught suction and is fitted with a 30 cm 8 W ultraviolet tube. The collection beaker has a volume of 500 ml and polyester netting around the light source prohibits moths and bigger insects from entering the trap.

Figure 2.1 Downdraught suction light trap used for collecting *C. imicola* from dawn until dusk. A UV light attracts the insects and fine netting prohibits larger insects and moths from entering the trap. The insects are collected in a collection beaker containing phosphate buffered saline (PBS). Image courtesy of Dr. G.J. Venter, ARC-OVI.

The insects were collected from dusk until dawn with the light traps being fitted 1,4 m above ground level at opposite ends of open sided barns. These barns housed around 20-40 cattle and the traps were hung as close to the cattle as practically possible. To prevent interference, the traps were stationed 15 m apart. The insects were collected in PBS and retrieved in the morning.

Insects of the genus *Culicoides* can be distinguished from one another based on a specific wing pattern. The method used for sub-sampling was followed as described by Van Ark and Meiswinkel (1992). The collected insects were placed on a chill table and observed using a stereomicroscope. Since only female *Culicoides* feed on blood, 99% of the insects caught in light traps around cattle were female. The females were distinguished from the males based on morphological characteristics (Figure 2.2). The abdominal pigmentation of the females that recently fed on blood was a deeper brown colour and bodies of the females tended to be rounder than those of their male counterparts (Dyce, 1969). The antennae of the females were also significantly less feathered than those of the males, enabling easy separation from the males.

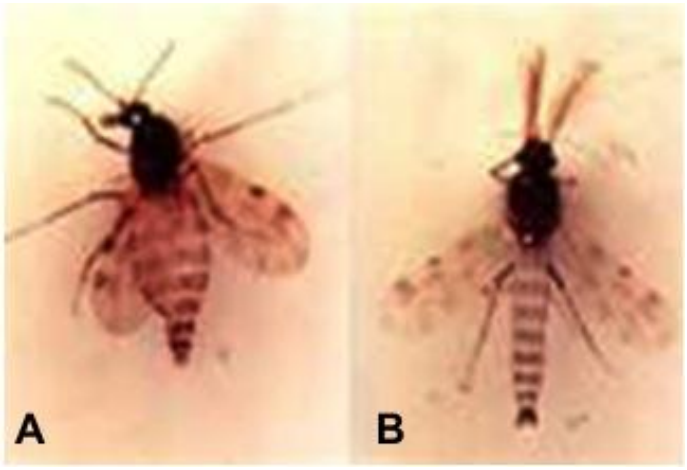

Figure 2.2 A *C. imicola* female (A) and male (B). Females are easily distinguished from the males based on morphological features such as body pigmentation, antennae and body shape. Image courtesy of Dr. G.J. Venter, ARC-OVI.

After sorting, around 1200 females were suspended in 5 ml of PBS, frozen at -20°C and couriered from ARC-OVI to the Pietermaritzburg campus of the University of KwaZulu-Natal where the remainder of the experimental work for this part of the study was conducted.

2.2.2 Preparation of a *C. imicola* protein homogenate

There are a variety of tools available for the homogenisation of insect samples. One of these techniques, which in previous studies was successful in homogenising samples containing smaller particles, is sonication. Sonication is a procedure by which samples are disrupted through the rapid expansion and contraction of pressure by a probe. This process occurs at high frequencies and the rapid oscillation of the current releases shock waves resulting in cell disruption.

Immediately upon receipt, the sample containing around 1200 midges were thawed on ice after which the midges were collected by centrifugation (10 000 *g*, 5 minutes, 4°C). The insect pellet was resuspended in 2 ml grinding buffer [125 mM Tris-HCl, 4% (w/v) sodium dodecyl sulphate (SDS), 20% (v/v) glycerol, pH 6.8] and subjected to short bursts of high intensity sonication using a Virtis VirSonic 60 sonicator (SP Scientific, Pensylvania, United States of America) set at an intensity of 10. Sonication was performed on ice for approximately 20 minutes. Since

sonication has the ability to generate a tremendous amount of heat, these short bursts were performed on ice to prevent denaturation of the proteins while homogenising the sample. Following sonication, the sample was centrifuged (16 000 *g*, 45 minutes 4°C). The supernatant of the homogenate was aliquoted into volumes of 100 µl and stored at -80°C until use. To prevent protein degradation, all steps were performed at 4°C.

2.2.3 SDS-PAGE

Routine analysis of purified protein samples can be evaluated under reducing or non-reducing conditions by means of polyacryamide gel electrophoresis (PAGE) (Laemmli, 1970; Sambrook & Russel, 2001). By boiling the protein in the presence of SDS, the interaction of the protein molecule with SDS masks the charge of the protein with the negative charge of dodecyl sulfate. This results in all proteins having the same anionic migration and charge to mass ratio. Furthermore, the treatment of the protein with β-mercaptoethanol will denature proteins into their constituent polypeptide chains, which will bind the SDS more effectively. The proteins can therefore be separated solely on size difference by migration through the polyacrylamide gel matrix. Finally, the proteins can be visualised by staining techniques such as Coomassie brilliant blue or silver staining.

Running gels with a final concentration of 12,5% acrylamide (Table 2.1) was prepared using Solution A [30% (w/v) acrylamide, 2,7% (w/v) bis-acrylamide], Solution B [1.5 M Tris-HCl, pH 8.8], Solution D [10% (w/v) SDS], and dH_2O. Polymerisation was catalysed through the addition of Solution E [10% (w/v) ammonium persulfate] and Tetramethylethylenediamine (TEMED). The 4% stacking gels' composition differed slightly in that Solution B was replaced with Solution C [500 mM Tris-HCl, pH 6.8]. The running gels were prepared by mixing its components in a beaker before the addition of solution E and TEMED.

Table 2.1 Running and stacking gel volumes used in the preparation of SDS-PAGE gel.

Reagent	Volume	
	Running Gel (12,5%)	**Stacking gel (4%)**
A	6,25 ml	940 µl
B	3,75 ml	-
C	-	1,75 ml
D	150 µl	70 µl
E	100 µl	35 µl
dH_2O	4,75 ml	4,3 ml
TEMED	10 µl	15 µl

The running gel solutions were cast between two glass plates (70 x 76 mm) of an assembled Bio-Rad Mini Protean® gel casting apparatus. The running gel was overlaid with dH_2O to allow for even polymerization for a period of 1,5 hours. The water was poured off after which the stacking gel was prepared and cast on top of the running gel. Ten-well combs were inserted and again, the gels were set aside to solidify for about an hour. The gels were used immediately.

Preparation of the protein samples for SDS-PAGE had equal volumes of *Culicoides* homogenate. The sample preparation was made up by adding 1 volume sample buffer (Appendix I) to 4 volumes protein sample to give a final volume of 20 µl. The samples were boiled at 98°C for 3 minutes unless otherwise stated. The sample was allowed to cool down to room temperature after which it was loaded onto the gel. A protein molecular size marker (Fermentas SM26620) was used to estimate protein size. The gel was electrophoresed in 1x Tris-Glycine SDS buffer (TGS Buffer) at a constant current of 80 V for 30 minutes after which the current was increased to 200 V until the pink dye of the marker neared the bottom of the gel. The gels were stained in staining solution [0.125% (w/v) Coomassie brilliant blue R-250, 50% (v/v) methanol, 10% (v/v) acetic acid] for 1 hour after which the gels were destained in several changes of destain [30% (v/v) methanol, 10% (v/v) acetic acid]. The stained gels were placed between two plastic sheets and digitised using an HP digital document scanner.

2.2.4 Gelatin-based substrate SDS-PAGE (Zymography)

Proteinase activity of proteins can be detected and characterised based on their ability to hydrolyse casein or gelatin co-polymerised within the running gel of polyacrylamide gels. This technique is more commonly known as zymography. Basically, protein samples are denatured during SDS-PAGE in the absence of a reducing agent such as β-mercaptoethanol. The samples are also not heated, as heating would interfere with the refolding of the enzyme. After electrophoresis, the SDS is removed from the gel and the enzymes renatured by washing the gel in an unbuffered Triton X-100 solution (Heussen & Dowdle, 1980). The gel is then incubated in a suitable digestion buffer at 37°C. Proteolytic activity was subsequently visualised following staining with amido black stain, as areas of gelatin digestion would appear as clear bands against a darkly stained background.

The technique for SDS-PAGE preparation was adapted from that described in section 2.2.3, in that 0.1% (w/v) gelatin was integrated with the gel to allow the detection of proteinases (Heussen & Dowdle, 1980). This was carried out by adding an additional 1,5 ml of 1% (w/v) gelatin in solution B (Section 2.2.3) to the remainder of solution B (2,25 ml). This was added to the rest of the solution for casting a 10% running gel (Table 2.2). The stacking gel used was identical to that described in Section 2.2.3. Non-reducing loading buffer was used (Appendix I) and the SDS-PAGE was carried out as described in Section 2.2.3.

Table 2.2 Running gel volumes used in the preparation of gelatin-based substrate SDS-PAGE gel.

Reagent	Volume
	Running Gel (10%)
A	5 ml
B	2,25 ml
D	150 µl
E	100 µl
dH_2O	5,9 ml
1% Gelatin	1,5 ml
TEMED	10 µl

Following electrophoresis, the gels were washed at room temperature in two changes of 2,5% (v/v) Triton X-100 over a period of 1 hour. The gels were incubated overnight at 37°C in incubation buffer [50 mM Tris-HCl, pH 8.0]. Staining was performed using stain solution [0.1% (w/v) amido black, 30% (v/v) methanol, 10% (v/v) acetic acid] over a period of 45 minutes after which the gels were destained in several changes of destain (Section 2.2.3). Gelatin digestion indicated proteolytic activity, which was visualised by the presence of clear bands in the gel. The gels were digitised as described in Section 2.2.3.

2.3 Results and Discussion

2.3.1 Preparation of a total protein extract from *C. imicola*

In order to extract the total amount of proteins present in *C. imicola*, all of the midges were homogenised mechanically to a pulp and the cells lysed with a detergent buffer. Nucleic acid material was released and the cell debri was removed by centrifugation. The homogenate supernatant was used for subsequent analyses.

The total protein homogenate of *C. imicola* was separated by SDS-PAGE. Even though the last centrifugation step removed a significant amount of cell debri, there were still some impurities present in the sample as is evident from the background staining observed in Figure 2.3. The molecular mass of the trypsin-like serine protease of *C. sonorensis* is around 29 kDa (Darpel *et al.*, 2011). Therefore, the protein of interest was also expected to have molecular mass of around 29 kDa. Since little is known about the proteins of *C. imicola*, different boiling times were used for denaturing the sample before separation on the SDS-PAGE. By boiling the sample in the presence of SDS, the tertiary structure of the proteins is denatured, and in the presence of a reducing agent the sites of the individual peptide chains can be determined. This results in the formation of a single band for separation based on the charge and molecular mass of the protein in the SDS-PAGE gel. To aid in the optimal denaturation of the proteins, the samples were

boiled at different periods of time, namely 2 minutes, 5 minutes and 10 minutes respectively. Following boiling the samples were centrifuged to collect the condensation that developed in the cap of the tube.

Four clear protein bands of around 50 kDa, 40 kDa, 29 kDa and 25 kDa were visible following boiling the samples for different periods of time (Figure 2.3, Lanes 2-4). A few fainter bands representing proteins of about 95 kDa, 80 kDa and 37 kDa were also visible. As seen in Figure 2.3, an increase in the duration of sample boiling resulted in the degradation of the proteins (Lanes 3-4), and subsequently a decrease in the intensity of the proteins. This was more prominent for the lower molecular mass proteins as proteins larger than 40 kDa were left unchanged with different boiling times. This could possibly be due to a weaker reaction of the larger proteins in the presence of the reducing agent at the high temperature (Sambrook & Russel, 2001).

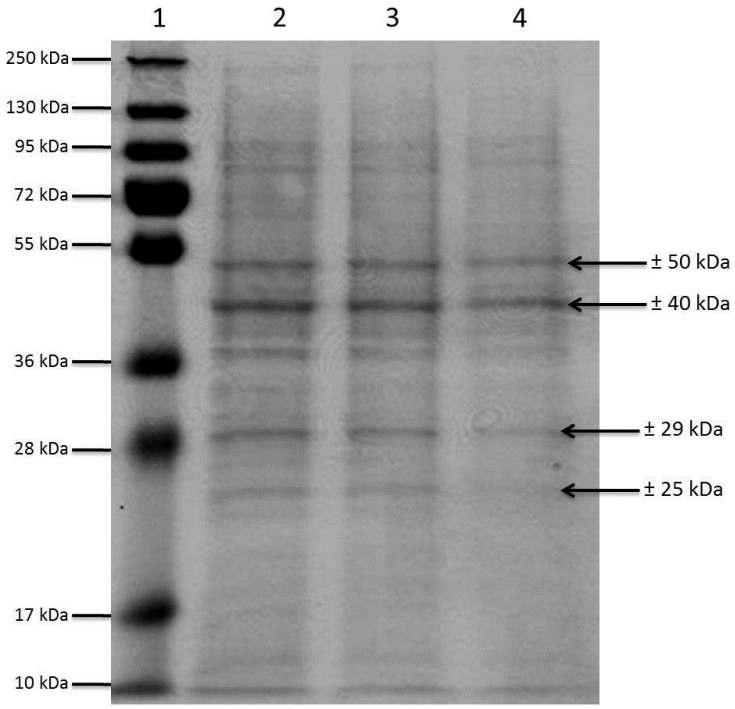

Figure 2.3 SDS-PAGE analysis of the total protein extract of around 1200 *C. imicola* midges. Lanes: 1) 6,5 µl PageRuler protein marker; 15 µl of *C. imicola* total protein extract boiled for 2) 2 minutes; 3) 5 minutes; 4) 10 minutes.

The proteins visible in lanes 2-4 at 25 kDa and 29 kDa were in the same molecular mass range as reported for the trypsin-like proteases in *C. sonorensis*. In order to further investigate these two proteins and to determine if they were trypsin-like serine proteases, proteolytic analyses under non-denaturing conditions were performed.

2.3.2 Analysing total protein extracts of *C. imicola* for proteolytic activity

To determine if the *C. imicola* proteins shown in Figure 2.3 have proteolytic activity, a very sensitive zymography assay was used. In this assay, gelatin was used as a substrate for the proteases and was incorporated into the SDS-PAGE gel. Proteins were separated as described in Section 2.2.4 under non-denaturing conditions. After electrophoresis the SDS was removed and the gel was incubated in an incubation buffer for three hours.

Viral proteases are constantly revealing new strategies in which they cleave the peptide bond to aid in the replication process (Darpel *et al*., 2011). New knowledge on the three dimensional structures (Manole *et al*., 2012), biological viral functions of proteases and the unique control mechanisms for proteolysis are constantly gained, giving much needed information on the ways in which posttranslational modifications occur during the replication of viruses (Babé, & Craik, 1997). Some of the separated proteins did have proteolytic activity as seen by the contrasting white bands against the darkly stained background (Figure 2.4, Lane 2). Distinct proteolytic activity was clearly visible at around 50 kDa, 40 kDa, 34 kDa, 28 kDa and 25 kDa. Of these five protease bands, four corresponded to the sizes of the proteins detected in Figure 2.3, Lane 2. There was no proteolytic activity at 29 kDa. However, there was proteolytic activity at 28 kDa. The precursor for trypsin is trypsinogen (Creighton *et al*., 1996), and cleavage of the precursor will activate trypsin. This mechanism of cleavage could possibly explain why activity was visualised at a smaller molecular mass than the corresponding protein. Therefore, the activity at 28 kDa may be due to cleavage of a precursor of the 29 kDa protein observed in Figure 2.3, Lane 2.

A smear of activity was visible from around 250 kDa up to about 72 kDa in Figure 2.4, Lane 2. This smear could be as a result of the sample not being pure and containing a large mixture of proteases (Wagner *et al.*, 2002). Within this smear, stronger proteolytic activity was visible at around 80 kDa. This activity could be related to the faint 80 kDa protein seen in Figure 2.3, Lane 2. Almost no proteins were visible at 34 kDa in Figure 2.3, Lanes 2-4, even though proteolytic activity was seen at 34 kDa of Figure 2.4, Lane 2. The concentration of this protein might have been so low that it resulted in it being invisible on SDS-PAGE. Gels where gelatin are incorporated as a substrate is known for being highly sensitive to protease activity. Based on this, it is possible that the activity of this protein could be visualised even though no protein was seen on the SDS-PAGE gel.

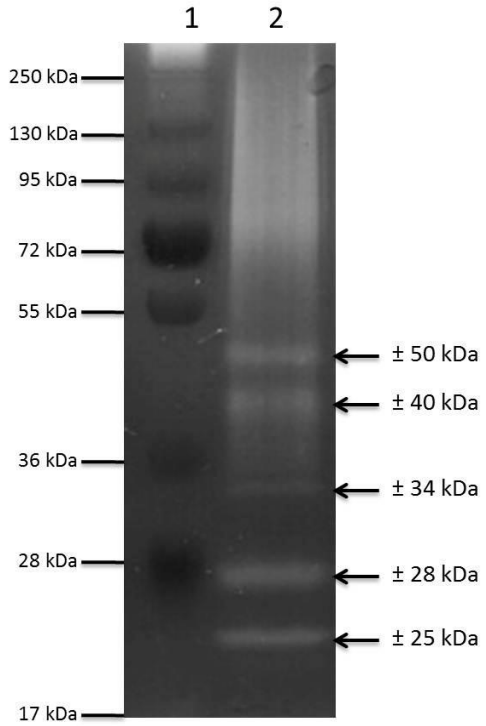

Figure 2.4 Gelatin-based substrate SDS-PAGE analysis of the proteolytic activity of proteins in a *C. imicola* homogenate. Lanes: 1) 6,5 µl PageRuler protein marker; 2) *C. imicola* homogenate supernatant.

To better identify these proteases and possibly characterise them, their activity had to be further investigated. This was done in the presence of different protease inhibitors.

2.3.3 Characterisation of *C. imicola* proteases using protease inhibitors

A protein of around 29 kDa was detected in Figure 2.3, Lane 2. This protein was subjected to gelatin SDS-PAGE that resulted in proteolytic activity being visualised at an estimated 28 kDa as seen in Figure 2.4, Lane 2. It is possible that the protein of Figure 2.3, Lane 2 is the same protein that shows proteolytic activity in Figure 2.4, Lane 2. However, this could only be speculated as the gelatin SDS-PAGE was conducted under non-denaturing conditions. The salivary protein identified in *C. sonorensis* was a trypsin-like serine protease (Darpel *et al*., 2011). To try and determine if the 28 kDa of *C. imicola* was also a trypsin-like serine protease containing a precursor, this protease had to be incubated in the presence of protease inhibitors.

Following electrophoresis and elimination of SDS through washing, the gels were incubated in incubation buffer containing different protease inhibitors. These inhibitors included antipain, a protease inhibitor that specifically targets the inhibition of trypsin, tosyl phenylalanyl chloromethyl ketone (TPCK) for the inhibition of chymotrypsin and ethelenediaminetetraacetic acid (EDTA), an inhibitor responsible for the inhibition of metalloproteases (Salvesen & Nagase, 2001).

When comparing Figure 2.5 A (Uninhibited control) with Figure 2.5 B, the proteolytic activity visible at 50 kDa in Figure 2.5 B, Lanes 2-3 was strongly inhibited by antipain and TPCK. Significant inhibition of this protease was also achieved by EDTA as low levels of activity was visible (Lane 4). Inhibition of the 50 kDa protease by three different classes of inhibitors probably means that a mixture of proteases were present in the sample (Wagner *et al.*, 2002) which may have resulted in proteolytic activity of different proteases at the same molecular mass. The incubation of *C. imicola* in the presence of antipain seemed to result in some inhibition of the 28 kDa protease and some form of inhibition of the 25 kDa protease (Lane 2). Even though very low activity was visible, it is still possible that further optimisation of the inhibitor concentration and inhibition time, might result in better inhibition. No inhibition of the 28 kDa protease was seen in the presence of TPCK and EDTA (Lanes 3-4).

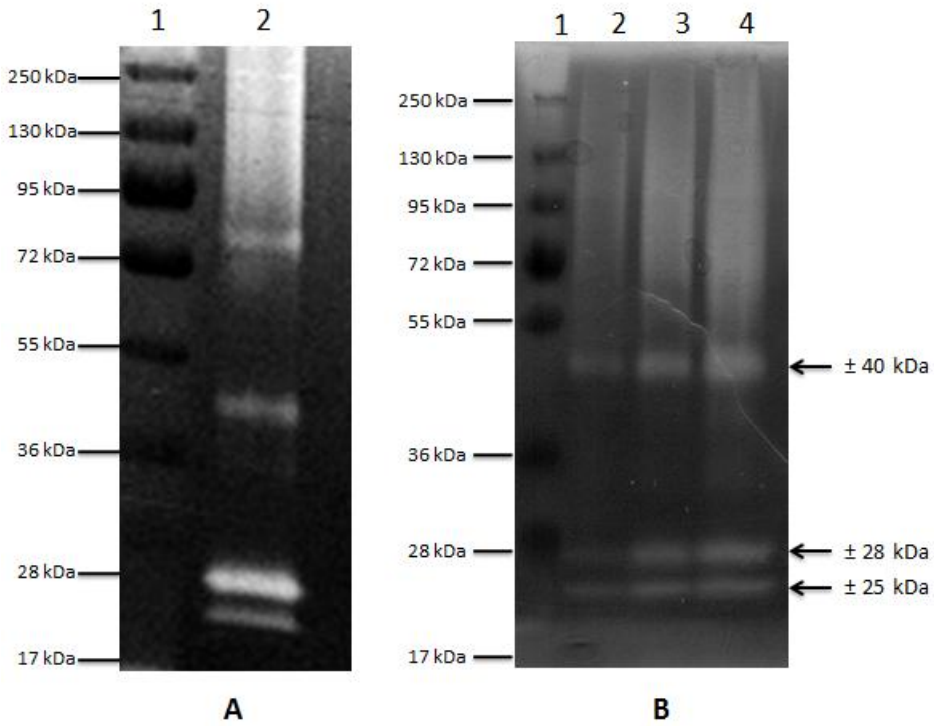

Figure 2.5 Gelatin-based substrate SDS-PAGE analysis of the effect of protease inhibitors on proteolytic activity of a *C. imicola* homogenate. Figure A lanes: 1) 6,5 µl PageRuler protein marker; 2) *C. imicola* homogenate supernatant. Figure B lanes: 1) 6,5 µl PageRuler protein marker; 2) Antipain, 1 µM; 3) TPCK, 100 µM; 4) EDTA, 5 mM.

Based on the results visualised in Figure 2.5, the possibility for this protease to be a metallo- or chymotrypsin-like protease was poor. Since antipain was the only inhibitor capable of inhibiting the 28 kDa protease, it can be concluded that this protease possibly has trypsin-like characteristics.

2.4 Summary

In a previous study performed on the vector for BTV, *C. sonorensis*, a 29 kDa trypsin-like proteinase that plays a role in cleavage of the structural proteins of BTV was discovered (Darpel *et al*., 2011). The vector *C. imicola* primarily transmits AHSV (Bouayoune *et al.*, 1998; Mellor *et al.,* 2000). Since *C. sonorensis* is also capable of transmitting AHSV (Boorman, 1974), it was hypothesised that *C. imicola* might have a similar protease to increase infectivity in AHSV. Hence, this part of the study had four main objectives, namely to obtain *C. imicola* through the

use of downdraught suction light traps and sorting of the insects based on subgenus and gender. The second objective was to extract all the proteins of the midges. This was done by homogenising the insects through sonication and separating the proteins by means of SDS-PAGE. The third objective was to confirm protease activity in the protein extract through the use of gelatin based substrate SDS-PAGE and finally, to try and determine if the activity was trypsin-like. The latter was done with the aid of catalytic class specific protease inhibitors incorporated in the gelatin SDS-PAGE incubation buffer.

The first objective was met with the identification of four major and three minor protein bands on the SDS-PAGE. Amongst these proteins was a 29 kDa protein. This protein band was of similar size to that of the salivary protein in *C. sonorensis* presenting trypsin-like characteristics (Darpel *et al*., 2011). An assay for detecting proteolytic activity was used to determine if any of these proteins were active proteases. The possibility of the proteins identified was confirmed when proteolytic activity was visualised at molecular masses very similar to the molecular mass of the proteins separated on the SDS-PAGE. Trypsinogen, the precursor for trypsin, is cleaved in order to activate trypsin (Voet & Voet, 1995). The 29 kDa protein showed activity at an estimated 28 kDa, possibly demonstrating that the protein is trypsinogen-like with the 28 kDa protease probably being trypsin-like. This hypothesis was strenghtened when the 28 kDa protease was inhibited by the trypsin inhibitor, antipain. The chymotrypsin and metalloprotease inhibitors, TPCK and EDTA did not inhibit the 28kDa protease. However, both inhibitors showed partial inhibition of the 40 kDa proteolytic activity.

The conclusion of the work described in this chapter is that a 29 kDa trypsin-like protease is present in *C. imicola*. This trypsin-like protease is also similar in size to that of the trypsin-like serine protease identified in *C. sonorensis* (Darpel *et al*., 2011). The next part of this study was to develop a bacterial expression system for the recombinant expression of *C. sonorensis* trypsin-like serine protease A.

CHAPTER 3

Bacterial expression of a 29 kDa *Culicoides sonorensis* trypsin-like protease

3.1 Introduction

In Chapter 2 it was established that *C. imicola* has a protein very similar in size to that of the 29 kDa salivary protease found in *C. sonorensis* and this particular *C. imicola* protein had trypsin-like protease activity. Few studies have been done on the proteases of *C. imicola* and therefore, no nucleic acid or amino acid sequencing data is available relating to these proteins. However, the 29 kDa proteases present in the saliva of *C. sonorensis* responsible for enhanced infectivity of AHSV and BTV has been sequenced and the results published (Darpel *et al*., 2011). Therefore, it was decided to try to clone and express this *C. sonorensis* trypsin-like serine protease A.

Bacterial expression systems, especially those using *Escherichia coli* (*E. coli*), are the preferred workhorses for expression due to rapid growth, high production yield and ease of manipulation of the bacteria (Graumann & Premstaller, 2006). The development of a bacterial expression system for recombinant proteases can contribute substantially to the investigation of these enzymes. Another advantage of these systems is the generation of large quantities of protein. This is especially helpful in cases where proteins occur in small amounts or are challenging to obtain (Terpe, 2006). The manipulation of proteins through modifications such as histidine tagged fusions can also aid in the purification of these proteins (Terpe, 2003).

One of the biggest shortcomings in the expression of recombinant proteases is the fact that their proteolytic activity and hydrophobicity could be toxic for the host cells (Mullins, 2010). Therefore, specific constructs with N-terminal fused tags need to be generated to protect the host cell from the proteolytic activity of the expressed proteases (Wan *et al*., 1995). Even though proteases have been expressed in

bacterial cells for long (Wan *et al*., 1995; Shi *et al*., 2009), no expression system is available for the bacterial expression of proteolytically active recombinant *C. sonorensis* trypsin-like serine protease A. Therefore, the first objective of this part of the study was to clone the open reading frame of late trypsin-like serine protease A into the cold expression vector pColdIII. The second objective was to express the late trypsin-like serine protease A with a C-terminal histidine tag in a bacterial host. The histidine tag should enable purification of the recombinant protein enabling a more detailed investigation of the protein (Terpe, 2003). The final objective was to verify proteolytic activity of the recombinant enzyme using zymography and to classify the expressed enzyme using protease inhibitors. This will aid in further studies of the specific proteins' functions. The long term goals of the work done in this part of the study is to develop a recombinant expression system that could be used for expressing proteolytically active *Culicoides* proteases responsible for the cleavage of VP2 in AHSV and BTV, resulting in increased infectivity.

To summarise, the main aims of this part of this study was to clone *C. sonorensis* late trypsin-like serine protease A into pColdIII, express and purify this recombinant enzyme using metal chelate affinity chromatography and finally proteolytically characterise this protease.

3.2 Materials and Methods

3.2.1 Source of the *C. sonorensis* late trypsin-like protease A coding sequence

The nucleotide sequence of the gene coding for the late trypsin-like serine protease A from C. *sonorensis* was published (Darpel *et al.*, 2011) and available at GenBank (Accession: AY603563.1). The open reading frame of the nucleotide consensus sequence was purchased from GenScript. Restriction sites at the 5’ and 3’ position, a serine-glycine linker and a C-terminal histidine tag was

incorporated into the sequence (Figure 3.1). The codon was optimised for expression in bacterial cells.

CATATGGGATCCATGCAGTTATTTCAAGTAATTTTCATTTTGATTACATTTTCCTTCATC
GAGATATTTGGTGATGAGGTTCAAGAAATTGTTCCAGTTCCACCTTTAAATTACACTAAA
CCACCTTACAGTGTAAAAATTGTTGGTGGTTCTCCTGCCCGCTTGCATCAATTTCCCTGG
CAAGCATCTATCACGTCATGCGAAGGAGGAAGTTGTTATATTTGTGGTGGTTCTTTGATC
TCGAAGCGATATGTTTTGACAGCAGCGCATTGCGCAGCCGGTTTAACCCGATTTGTCATT
GGTCTTGGTTCCAATTCGAGAAATCGACCTGCTGTTACGTTAACTTCAAATATAAAAGTT
GTACACCCACAATACGATGCTAAAAGCTTAGGAAATGATGTTGCTGTCATTAAATTACCC
TGGAGTGTGAAATTAAACAAGGCAATTCAACCAATCATTTTGCCACGCTCAAATAATACA
TACGATAATGCAAATGCAACTGTATCGGGATATGGGAAAACAAGCGCTTGGAGTTCGTCA
TCTGATCAATTGAATTTTGTTGATATGCGTATTATATCAAATGGTCAATGTAGAGAAATA
TTTGGATCAGTAATTCGTGATTCGAGTTTGTGTGCTGTAGGAAAAAATTTGTCAAGGCAA
AATGTATGTCAAGGTGATTCTGGAGGTCCATTGGTGGTTAAAGAAGGAAATTCAACTGTC
CAAGTCGGAGTTGTATCTTTTGTATCTGCAGCTGGATGTGCAGCTGGTTATCCATCTGGA
TATGCTAGAGTTAGCTCATTTTATGAATGGATTGCCAATATGACAGACATTGATTTGAGC
AGCGGCGGTGGCGGCAGCCATCATCATCATCACCATTAAGTCGACGCGGCCGC

Figure 3.1 Nucleotide sequence of the bacterial codon optimised open reading frame of the late trypsin-like serine protease A. The 5' restriction sites are NdeI in orange and BamHI in pink, the nuceolotide open reading frame in black, serine-glycine coding linker in yellow, histidine tag in green and the 3' restriction sites SalI in purple and BamHI in blue. The start and stop codons are indicated in red.

The *Culicoides sonorensis* late trypsin nucleotide sequence was named *CulsonLTRYP* and will therefore be referred to as such in the text. *CulsonLTRYP* was cloned between the restriction sites NdeI and EcoRV in the multiple cloning site of the pUC57 simple cloning vector. The purchased sequence was received in lyophilised form and dissolved with nuclease free H_2O to a final concentration of 4 µg/µl.

3.2.2 Cell lines and expression vector used in this study

Two different strains of *E. coli* were used in this study, namely JM109 and Origami cells. *E.coli* JM109 is an all-purpose strain commonly used for cloning purposes (Sambrook & Russel, 2001). The JM109 cell line carries the *rec*A1 and *end*A1 mutations contributing to plasmid stability and high quality plasmid deoxyribonucleic acid (DNA) preparation. Origami cells are an expression cell line

commonly used in protein expression studies (Hunt, 2005). A mutation of the *trx*B and *gor* genes within the Origami cell genes create an increased oxidising environment within the cell. This allows for disulfide formation within the cytoplasm, which contributes to an increase in the activity of the protein (Bessette *et al.*, 1999). A member of our research group generously provided the chemically competent Origami and JM109 cells used in this study.

The vector used for expression in this study was the cold-shock expression vector pColdIII (Figure 3.2). This vector contains the *csp*A cold-shock promotor, which is utilised for the efficient expression of the protein at low temperatures. Located downstream of *csp*A promotor is the 5' untranslated region, translation enhancing element and multiple cloning site. The *lac* operator is also situated downstream of the *csp*A promotor enabling strict control of expression.

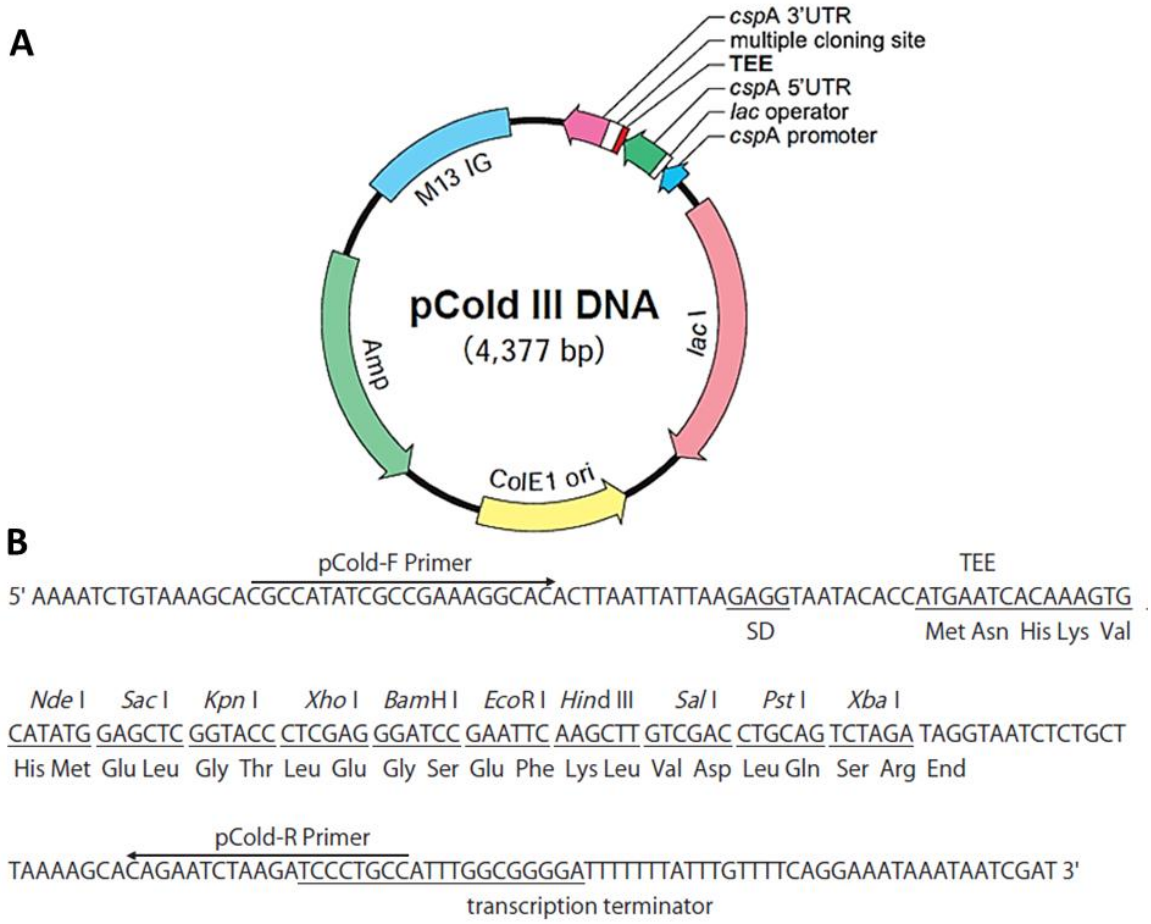

Figure 3.2 Plasmid map (A) and multiple cloning site (B) of the expression vector, pColdIII. Figure A illustrates the *csp*A promotor, *lac* operator *csp*A, 5' untranslated region, the translation enhancing element, multiple cloning site and the *csp*A 3' untranslated region. The green arrow indicates the ampicillin resistant gene (Taken from Takara cold shock expression system manual).

3.2.3 Preparation of electrocompetent JM109 *E. coli* cells

Electroporation is a method in which the electrical conductivity and permeability of the cell plasma membrane is increased with an external electrical field. This technique enables more effective transformation of plasmid DNA into bacteria (Neumann *et al.*, 1982). The technique of preparing electro-competent cells is much simpler than the preparation of chemically competent cells. In brief, the cells are grown to mid log phase, salts are removed with several washes of 10% (v/v) glycerol, DNA is mixed with the cells, placed in a chilled electroporation cuvette and an electrical pulse is applied (Sambrook & Russel, 2001).

A 5 ml culture of Luria Broth (LB) medium [1% (w/v) tryptone, 0,5% (w/v) yeast extract, 1% (w/v) NaCl] (Sambrook & Russell, 2001) was inoculated with a single *E. coli* colony, which was streaked from a glycerol stock. The culture was grown at 37°C overnight in the absence of antibiotics. Two 50 ml cultures of LB medium were incubated with 2,5 ml overnight culture and grown at 37°C with agitation until an optical density (OD) of 0,4 to 0,6 at 600 nm was reached. The cells were immediately chilled on ice for 30 minutes and divided into four pre-chilled 50 ml conical tubes. From this point forwards, all steps were performed at 4°C. The cells were harvested by centrifugation (2000 *g*, 15 minutes, 4°C). The supernatant was discarded and each of the pellets was resuspended in 25 ml ice-cold sterile deionised water. Centrifugation was repeated after which the pellets were resuspended in 5 ml ice-cold 10% (v/v) glycerol and again centrifuged (3000 *g*, 20 minutes, 4°C). This step was repeated twice. Upon completion of centrifugation, the cells of all four tubes were resuspended in 5 ml ice-cold 10% (v/v) glycerol and pooled. The purity of the cells was tested by adding 60 µl of cells to a pre-chilled electroporation cuvette and applying a pulse of 1.8 kV for 1 ms. If arching occurred, it was an indication that additional washing of the cells with glycerol were needed. The cells were dispensed in aliquots of 50 µl in pre-chilled tubes and snap-frozen in a container of liquid nitrogen. The cells were stored at -80°C.

3.2.4 Transformation of electrocompetent JM109 *E. coli* cells

All steps in the transformation of electrocompetent *E.coli* cells were performed as described in the literature (Sambrook & Russel, 2001). The DNA of pUC57*CulsonLTRYP* and pColdIII to be used in transformation was diluted to 10 ng/ml. A Bio-Rad GenePulser Xcell electroporator and cuvettes were used.

An electroporation cuvette was placed at -20°C, 20 minutes prior to the start of transformation. A 50 µl aliquot of frozen electrocompetent cells was removed from storage and thawed on ice. The cells were gently mixed with 3 µl of DNA and incubated on ice for 45 minutes. Taking care that no air bubbles formed, the cell slurry was transferred to the pre-chilled electroporation cuvette. A pulse of 1.8 kV was applied for 1 ms. The cells were allowed to recover for 1 hour at 37°C with gentle agitation by adding 1 ml super optimal broth with catabolite repression

culture (SOC) [2% (w/v) tryptone, 0,5% (w/v) yeast extract, 0.05% (w/v) NaCl, 250mM KCl pH 7.0, 2 M $MgCl_2$ and 1 M glucose] (Sambrook & Russell, 2001) to the cells immediately after pulsing. This allowed the cells to express the antibiotic resistance marker encoded by the plasmid before selection when antibiotics are applied. A total volume of 100 µl transformed cells of this culture was spread out on LB agar plates containing 50 µg/ml ampicillin. The plates were incubated at room temperature to absorb the liquid after which they were inverted and incubated at 37°C overnight.

3.2.5 Mini-preparation of plasmid DNA

With mini-preps, pure plasmid DNA can be isolated using only 1-2 ml of culture medium. This makes the mini-prep method a very cost effective way for isolating small quantities of DNA. Two times 10 ml LB culture medium samples, containing 50 µg/ml ampicillin, were respectively inoculated with a single colony of pUC57*CulsonLTRYP* and pColdIII from the overnight streaked out plates and incubated at 37°C with vigorous shaking. A volume of 1,5 ml of each of the overnight cultures were transferred to 1,5 ml microfuge tubes and centrifuged (11 000 *g*, 30 seconds).

The pUC57*CulsonLTRYP* and pColdIII plasmids were purified using a commercial plasmid purification kit, the Nucleospin® plasmid DNA Miniprep kit system (Macherey-Nagel, Germany). The kit was designed with the aim of isolating high-quality plasmid DNA by using a silica membrane column. These columns isolate the plasmid DNA, with the incorporation of a ethanol wash. This wash effects the elimination of substantial amounts of genomic DNA, cell debri and precipitated proteins from the purified plasmid DNA (Birnboim & Doly, 1979). The protocol for purifying the DNA of transformed competent bacterial cells by means of centrifugation was followed as described by the manufacturer. There are four major steps in the procedure of isolating plasmid DNA using the Nucleospin® kit. These steps include the preparation and lysis of the bacterial cell walls, purifying the DNA by means of centrifugation, washing of the column and elution of the plasmid DNA. By incubating the bacterial cells for the set time, the two main components of the cell lysis solution namely, sodium hydroxide (NaOH) and SDS

denature the proteins within the cell to ensure the release of plasmid DNA without over-exposing the plasmid DNA to denaturing conditions. The denatured components are precipitated by the addition of the neutralization solution. This step creates appropriate conditions for the binding of the plasmid DNA to the silica membrane. Contaminants such as endonucleases, ribonucleic acid (RNA) and endotoxin proteins are not capable of binding to the column and will pass through the column upon washing (Birnboim & Doly, 1979). Following the washing of the column with the ethanol wash buffer, the plasmid DNA was eluted with 30 µl nuclease free water. Four preparations were performed per plasmid and like samples were pooled.

3.2.6 Spectrophotometric quantification of isolated DNA

A pure DNA solution with an OD of 1 at 260 nm (OD_{260}) corresponds to 50 ng/µl (Sambrook & Russell, 2001). The NanoDrop® (Thermo Scientific, Wilmington, United States of America) instrument was used for the spectrophotometric quantification of the isolated DNA. The purity of a DNA solution can be established by the A_{260}/A_{280} ratio. Tryptophan and tyrosine residues are primarily responsible for the strong absorption at 280 nm by proteins. DNA that is pure is considered to have an A_{260}/A_{280} ratio of an observed value less than 1.8. Isolated DNA's concentration can be calculated by using an equation (Equation 3.1). However, the concentration of the pUC57*CulsonLTRYP* and pColdIII DNA was calculated automatically by the NanoDrop® instrument using the same parameters as those given in Equation 3.1. The instrument was always blanked with 18 Ω water, or an appropriate buffer for the sample to be analysed.

Equation 3.1 Calculation of DNA concentration (Sambrook & Russell, 2001)

$$[\text{dsDNA}] = A_{260} \times (50\ \text{ng/µl} \times \text{Dilution Factor})$$

3.2.7 Restriction endonuclease digestions

Restriction enzymes are used to treat the cloning vector in order to cleave the DNA at the site where the foreign DNA will be inserted. This is typically done by treating the vector DNA and the foreign DNA of interest with the same restriction enzyme, for example SmaI (Sambrook & Russell, 2001).

Restriction enzymes purchased from Fermentas were used according to the manufacturers' recommendations for the digestion of plasmids. Reaction mixtures contained 500 ng plasmid DNA, with the appropriate buffer (Table 3.1) for the enzyme used. The buffers were made up to a final concentration of 1x. Reaction mixtures contained 20 units of enzyme per microgram DNA. All restriction digestion reactions were performed overnight at 37°C. The reactions were finally analysed by means of agarose gel electrophoresis.

Table 3.1 Restriction enzymes and buffers used

Restriction enzyme	Recognition sequence	Buffer used
SalI	GTC GAC	Buffer O
NdeI	CAT ATG	Buffer O

3.2.8 Agarose gel electrophoresis

The analysis of restriction digestions and PCR reactions was performed by means of agarose gel electrophoresis. Standard procedures as described in the literature were followed (Sambrook & Russell, 2001). Unless otherwise stated, 1% (w/v) agarose gels (6 x 10 x 0.5cm) were prepared for all reactions by using 1x Tris-acetate-EDTA (TAE) buffer. The visualisation of the DNA fragments on an ultraviolet light trans-illuminator was made possible by adding 0.5 µg/ml ethidium bromide to the gel mixture.

Generally, 15 µl samples were mixed with 3,5 µl loading dye (Fermentas R0631) before loading onto the gels. Molecular weight markers (Fermentas SM1173) were always loaded in one of the lanes for the estimation of the DNA sizes.

Electrophoresis was conducted for 1 hour at 80 V using a Bio-Rad PowerPac basic system (Bio-Rad, Hertfordshire, United Kingdom). Directly following gel electrophoresis, the gels were photographed using the Syngene ChemiGenius Bio-Imaging system and accompanying GeneSnap software (Syngene, Cambridge, United Kingdom).

3.2.9 Gel purification of digested pUC57*CulsonLTRYP* and pColdIII DNA

One of the biggest advantages of using agarose gels is the availability of low melting temperature agarose. As the name suggests, DNA samples that were separated in a gel can be excised from the gel and heating can easily reliquefy these gels, returning the DNA to solution and the DNA can be recovered (Wilson & Walker, 2010). Throughout the whole study low melting point agarose gel was used.

The NucleoSpin® Extract II DNA extraction kit was used. All steps were performed according to the manufacturers' recommendations. Following restriction digestion, the DNA fragments to be used for cloning purposes were excised from the agarose gel. A buffer (NT) was added to the gel, which facilitated the dissolution of agarose at 50°C. To facilitate the separation of DNA from the agarose gel, the solution was spun through a column containing a silica membrane. This membrane allowed the binding of DNA in the presence of the chaotrophic salts in the NT buffer (Vogelstein & Gillespie, 1979). Contaminants such as soluble macromolecular components, ethidium bromide and salts do not bind to the silica membrane and were easily removed by washing the column with the buffer NT3 containing 70% ethanol. This lowered the ionic strength conditions of the membrane and enabled the elution of pure DNA. The DNA was eluted in a volume of 25 µl with nuclease-free water instead of buffer NE as described by the manufacturer.

3.2.10 Ligation reactions

An important step in molecular cloning is the creation of a recombinant plasmid with the introduction of the recombinant DNA into the host organism. This can be achieved by combining the prepared vector DNA with the foreign DNA. Furthermore, the addition of DNA ligase to the DNA mixture will result in the covalently joining of the vector and foreign DNA ends, ultimately creating a recombinant plasmid (Sambrook & Russell, 2001).

Ligations of the digested *CulsonLTRYP* (Insert) and pColdIII (Plasmid vector) to create a recombinant pColdIII*CulsonLTRYP* (Figure 3.3) were performed as described in the literature (Sambrook & Russel, 2001). The formula according to Sambrook and Russell (2001) depicted in equation 3.2 was used to determine the amount of DNA to be used for the ligation reaction. The ligation reaction mixtures consisted of a 3:1 vector to insert molar ratio (50 ng vector DNA: 150 ng insert DNA), 5 Weiss units of T4 DNA ligase and 1x ligation buffer.

Equation 3.2 DNA Calculation for ligation (Sambrook & Russell, 2001)

$$\text{Insert (ng)} = \frac{\text{Vector concentration (ng)} \times \text{Insert size (kDa)} \times 3}{\text{Vector size (kDa)} \times 1}$$

The ligation reaction was set up as follows: to prevent the formation of unwanted self-annealing of DNA, the plasmid and insert DNA were added to a 250 µl tube and heated at 65°C for 1 minute. This was followed by a slow cool down of the mixture at a rate of 1°C per second to 4°C on a Bio-Rad thermal cycler (Bio-Rad, California, United States of America). Annealing of the cohesive DNA ends were facilitated by the slow cool down step. DNA concentrations are high at this point since only plasmid and insert DNA are present. These high DNA concentrations will therefore result in inter-molecular annealing as opposed to intra-molecular annealing (Sambrook & Russel, 2001). T4 ligase and buffer were added after this step and the final volume made up to 20 µl. The ligation reaction was incubated at 4°C for 48 hours.

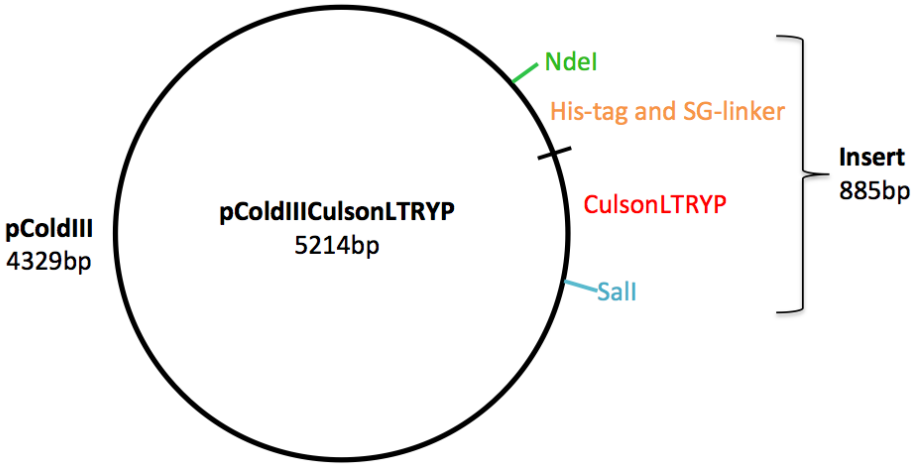

Figure 3.3 Plasmid map after *CulsonLTRYP* is cloned into the pColdIII expression vector. The 885 bp insert fragment consists of *CulsonLTRYP* (Red) with a fused C-terminal His-tag and SG-linker (Orange). The insert was cloned into the expression vector between the restriction sites SalI (Blue) and NdeI (Green).

3.2.11 Transformation of chemical competent Origami *E. coli* cells

Transformation of chemical competent *E. coli* was performed as described by Sambrook & Russel (2001). Chemical competent Origami cells previously prepared by another member our research team was generously provided for this part of the study.

Frozen chemical competent Origami cells of 75 µl aliquots were removed from storage at -80°C and thawed on ice. A volume of 10 µl ligation mixture containing the plasmid DNA was added to 75 µl competent cells and gently swirled to ensure thorough mixing of the contents. The cell mixture was incubated on ice for 60 minutes, followed by heat shock for precisely 90 seconds in a pre-heated circulating water bath set at 42°C. Immediately following the heat shock step, the cells were transferred to an ice bath for 30 minutes. A volume of 1 ml chilled SOC medium was added to the tubes, inverted a few times and incubated (185 rpm, 60 minutes, 37°C). Volumes of 150 µl transformed cells were spread out on LB agar plates containing 50 µg/ml ampicillin. In order for the liquid to absorb, the plates were incubated at room temperature for 15 minutes after which they were inverted and incubated at 37°C overnight.

3.2.12 Screening of recombinant pColdIII*CulsonLTRYP*

Colony screening of transformed bacteria

To determine if the size of the DNA cloned into the plasmid is correct, PCR can be performed from a single colony on the transformation plate. This serves as a rapid initial form of screening before performing plasmid isolation. Not only can designing primers specific to the sequence of the insert verify the presence of the insert, but also the orientation of the insert.

Using the primers *CulsonLTRYP*_For and *CulsonLTRYP*_Rev (Table 3.2), the coding sequence of *CulsonLTRYP* was PCR amplified. The PCR reaction mixtures contained 10 nmol each of dNTP, 25 pmol of each primer, 1x Takara ExTaq buffer and 2 units of Takara ExTaq polymerase in a final volume of 25 µl. Single, well-isolated colonies were picked from the transformation plate and washed into the PCR reaction mixture. For short-term preservation, the selected colonies were streaked onto a master plate. Positive colonies could be selected form the master plate containing ampicillin for inoculation.

Thermal cycling conditions were 5 minutes at 94°C, then 40 cycles of 94°C for 40 seconds, 55°C for 60 seconds and 72°C for 45 seconds. This was followed by a final extension of 10 minutes at 72°C. Thermal cycling was performed on a Bio-Rad thermal cycler.

Table 3.2 *CulsonLTRYP* oligonucleotide primers used

Primer name	Oligonucleotide sequence (5' → 3')	Length (bp)	T_m (°C)
*CulsonLTRYP*_For	GGGATCCATGCAGTTATTTCAAG	23	54.1
*CulsonLTRYP*_Rev	CGCGTCGACTTAATGGTGATG	21	56.2

Restriction endonuclease analysis

A total of 3 positive colonies of the transformed pColdIII*CulsonLTRYP* were selected from the reference plate for restriction digestion. These colonies were

used to inoculate 3 ml LB medium containing 50 µg/ml ampicillin. The cultures were incubated overnight (185 rpm, 37°C).

Plasmid DNA was isolated from 1,5 ml of culture as follows: the cells were harvested by centrifugation (16 000 *g*, 5 minutes). The supernatant was discarded, and the cellular pellet was resuspended by adding 250 µl STET buffer [8% (m/v) sucrose, 5% (v/v) Triton-X100, 50 mM EDTA, 50 mM Tris-HCl buffer pH 8]. Lysozyme (1 mg/ml) was added to the STET buffer to lyse the bacterial cell wall. The mixture was boiled at 98°C for 1 minute. Immediately following boiling, the lysates were centrifuged (16 000 *g*, 8 minutes). The pellet was removed with a toothpick and 3 µl of a 20 µl/ml RNase A solution was added. The mixture was incubated at room temperature for about 10 minutes after which 250 µl isopropanol was added. Contaminating salts are dissoluted in the presence of isopropanol while DNA are retained as a precipitate. The mixture was centrifuged (16 000 *g*, 8 minutes). The supernatant was aspirated and the DNA washed with 600 µl 70% ethanol. The DNA was dried in a vacuum dryer (68°C, 10 minutes) and resuspended in 20 µl nuclease-free water. Of this plasmid preparation, 3 µl was digested using restriction enzymes as described in Section 3.2.7 and the DNA fragments were analyzed on a 1% (w/v) agarose gel.

DNA sequence determination

The recombinant plasmid DNA was isolated according to the method described in Section 3.2.5 and the DNA spectrophotometrically quantified as discussed in Section 3.2.6. Sanger sequencing was used to confirm that the recombinant plasmid contained the gene of interest. Samples were sequenced (DNA sequencing laboratory, Central Analytical Facility, University of Stellenbosch). The DNA sequence electropherograms were analyzed using FinchTV version 1.40 (www.geospiza.com/finchtv) and ClustalX2 (http://www.clustal.org/clustal2) was used to align the sequences obtained to the reference sequence.

3.2.13 Expression of *C. sonorensis* from pColdIII expression vector

Cloning of the DNA sequence of the protein of interest into a high-copy number plasmid that contains the *lac* promoter will enable expression of the protein of interest by the addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) (Stark, 1987). IPTG is a lactose analog and activates the *csp*A promoter, resulting in protein expression by the bacteria (Oehler *et al*., 1990).

For expression, the transformants were inoculated in 250 ml LB culture medium containing 50 µg/ml ampicillin and incubated overnight (200 rpm, 37°C). The overnight culture was aliquoted in 50 ml volumes and the cells were harvested by centrifugation (2 000 *g*, 20 minutes). The cell pellets were resuspended in 200 ml fresh LB medium containing 50 µg/ml ampicillin and incubated (150 rpm, 3 hours, 15°C). By reducing the culture temperature, the growth of *E. coli* was momentarily stopped, and the "cold-shock promotor" was specifically induced. Expression of *CulsonLTRYP* was then induced by addition of IPTG to the appropriate final concentration and continued incubation (150 rpm, 24 hours, 15°C).

3.2.14 Cell lysis using the BugBuster™ protein extraction reagent

A simple alternative to mechanical methods such as sonication for the release of expressed target proteins is the BugBuster™ protein extraction reagent. The non-ionic detergent in the formula gently disrupts the cell wall of *E. coli*, resulting in the liberation of soluble proteins free from the denaturing agent.

Cells were harvested from the liquid culture by centrifugation (10 000 *g*, 30 minutes, 15°C). Bugbuster™ protein extraction reagent containing Benzonase was added to the cell pellet according the manufacturers' instructions (5 ml per 1 g wet cell paste). The cells were resuspended by gentle vortexing. Cell lysis was allowed by incubation at room temperature for 20 minutes on a rocking platform. After incubation, 1 ml of the sample, which contained the total cell debri, soluble proteins and insoluble proteins was aliquoted and labeled "total protein fraction". The insoluble proteins and cell debri was removed by centrifugation (16 000 *g*, 30 minutes, 4°C). Following centrifugation, 1 ml of the cleared cell lysate was

transferred to a clean tube and labeled "soluble protein fraction". The remainder of the cleared cell lysate was used in further protein purification steps.

3.2.15 Metal chelate affinity purification of histidine tagged protein and ultra-filtration

The Protino® Ni-TED 2000 kit was used for affinity purification of histidine tagged proteins according to the manufacturers' guidelines. Recombinant polyhistidine-tagged proteins are purified by immobilised metal chelate affinity chromatography. Proteins bind based on the interaction between the polyhistidine tag of the recombinant protein and the strong pentadentate metal chelator and immobilised Ni^{2+} ions.

The Protino® Ni-TED columns were equilibrated with buffer LEW after which the cleared cell lysate containing the expressed protein of interest was passed through using gravity. The columns were washed 4 times with buffer LEW and the histidine-tagged proteins were eluted using 3 ml elution buffer (EB). All purification steps were performed at 4°C.

Ultrafiltration spin columns were used to concentrate the eluate fractions. The Vivaspin 6 ultrafiltration system was used. The eluate was transferred to the filtration device and centrifuged (3 000 *g*, 35 minutes, 4°C). The sample was concentrated from 3 ml to approximately 200 µl after which the concentrate was transferred to storage at 4°C. All concentration steps were performed at 4°C.

3.2.16 Silver staining of proteins

The silver staining technique was used for the visualisation of the expressed protein after SDS-PAGE. Silver staining is a very sensitive technique used for the visualisation of very small amounts of protein (300-500 ng). This technique has a considerably higher degree of sensitivity than that obtained with Coomassie brilliant blue R-250 staining (Blum *et al.*, 1987), whilst using very cheap equipment and materials. The basic principle of silver staining relies on the binding of proteins to silver ions. The ionic silver in solution can in return be reduced to its insoluble

metallic form to finally build up a detectable image made of finely separated silver metal.

All steps were carried out at room temperature on a shaking platform (35 rpm). To minimise the occurrence of background staining, glassware were meticulously cleaned. Following electrophoresis, the gels were soaked overnight in 100 ml fixing solution [50% (v/v), methanol, 12% (v/v) acetic acid, 0.5% (v/v) 37% formaldehyde] followed by incubating (3 x 20 minutes) in washing solution [50% (v/v) methanol]. The gel was soaked in pretreatment solution [4 mg/ml $Na_2S_2O_3.5H_2O$] for 1 minute, rinsed in dH_2O (3 x 20 seconds) and soaked in impregnation solution [0.2% (m/v) $AgNO_3$, 0.75% (v/v) 37% formaldehyde] for 20 minutes. Following rinsing in dH_2O (3 x 20 seconds), the gels were immersed in developing solution [60 g/l Na_2CO_3, 0.5% (v/v) 37% formaldehyde, 4 mg/ml $Na_2S_2O_3.5H_2O$]. As soon as the first protein bands became visible, developing solution was replaced with dH_2O until development was sufficient. Development was stopped by soaking the gels in stopping solution [50% (v/v) methanol, 12% (v/v) acetic acid] for 10 minutes. The gels were digitised as described in Section 3.2.3.

3.2.17 Fluorogenic peptide specificity and inhibitor profile assay

Proteolytic enzymes cleave the peptide bond, known as the scissile bond, of proteins. The protease-catalised hydrolysis of the scissile bond would result in the formation of two separate polypeptides with new carboxyl and amino termini. The Schechter and Berger notation is used to better understand the mechanism and specificity of this proteolytic action (Figure 3.4). The residues of the substrate on the N-terminal side of the scissile bond are denoted as P1, P2, P3..., whereas, residues on the C-terminal side of the scissile bond, are denoted as P1', P2', P3' etc. from the N to C direction. The specificity of proteases are based on their ability to recognise the substrate residues in the P1, P2, P3,..., and P1', P2', P3'..., positions (Smooker *et al.*, 2010; Cronk, 2012).

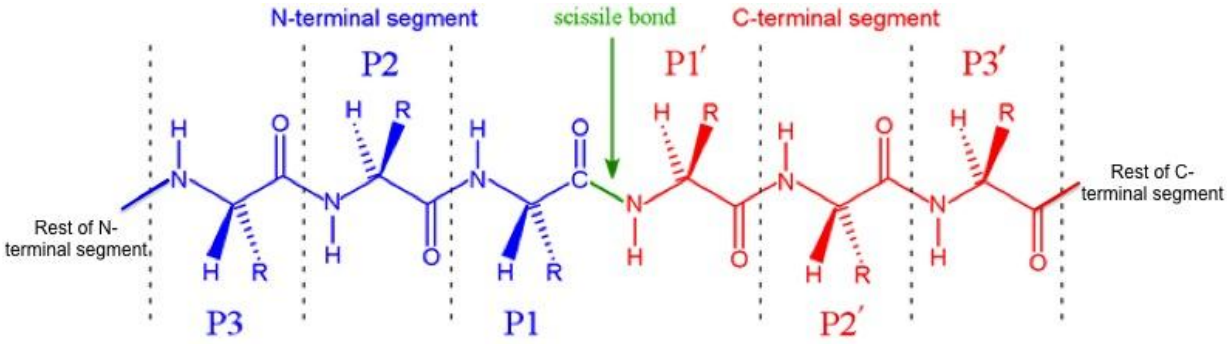

Figure 3.4 Schematic diagram for discussing the petidase active site according to Schechter and Berger notation. The residues of the substrate on the N-terminal side of the scissile bond (Green) are denoted as P1, P2, P3 etc. (Blue) in the C to N direction. On the C-terminal side of the scissile bond, substrate residues are denoted as P1', P2', P3' etc. in the N to C direction (Red) (Smooker *et al.*, 2010; Cronk, 2012).

Enzymatic activity can be measured by the use of peptide substrates containing highly fluorescent 7-amino-4-methylcoumarin (AMC) (Figure 3.5). The following substrates, Benzoyl-L-Arginine-AMC (BZ-L-Arg-AMC), Carbobenzoxy-L-Alanine-L-Arginine-L-Arginine-AMC (BOC-Ala-Arg-Arg-AMC), t-Butyloxycarbonyl-β-Benzyl-L-Aspartyl-L-Prolyl-L-Arginine-AMC (BOC-Asp(OBzl)-Pro-Arg-AMC), t-Butyloxycarbonyl-L-Valyl-L-Prolyl-L-Arginine-AMC (BOC-Val-Pro-Arg-AMC) and Benzyloxycarbonyl-L-Pyroglutamyl-Glycyl-L-Arginine-AMC (Z-Pyr-Gly-Arg-AMC) were used.

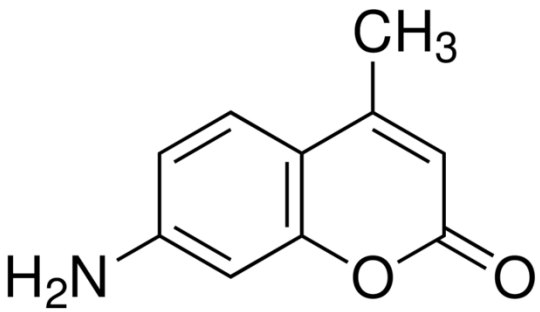

Figure 3.5 Structure of 7-Amino-4-MethylCoumarin (AMC). An amide bond is formed between the amino group of AMC and the carboxyl group of the P1 residue. This enables the measurement of peptidase activity capable of cleaving the amide bond, resulting in an increase in fluorescence.

Substrate stock solutions (1 mM) were prepared to a final volume of 1,5 ml. All reagents were preincubated at 37°C and assays carried out in FluorNunc® 96-well

fluorometry plates. In triplicate, assay mixtures contained r*CulsonLTRYP* (20 µl), 0.1% (w/v) Brij 35 (30 µl) and 25 µl assay buffer [200mM Tris-HCl, pH 8, 10 mM Dithiothreitol (DTT), 0.02% (m/v) NaN_3]. The reaction mixtures were activated by incubation (37°C, 15 minutes). A volume of 25 µl substrate solution (1 mM) was added to the reaction mixture and the fluorescence (Excitation 380 nm, emission 460 nm) was measured immediately. A FLUOstar micro plate fluorometer (Optima, Germany) and accompanying OPTIMA software was used for the analysis. Omitting r*CulsonLTRYP* from the reaction mixture served as negative controls.

Protease inhibitors can be classified as compounds that decrease the measured rate of enzyme catalysed hydrolysis of a particular substrate (Rawlings *et al.*, 2004). Enzyme inhibitors present themselves in many forms, such as being structurally very similar or the manner in which their reaction mechanisms occur (Salvesen & Nagase, 2001). The recombinant protease, r*CulsonLTRYP*, was incubated in the presence of a variety of protease inhibitors (Table 3.3) in order to classify the enzyme. All steps were performed exactly as described for the determination of peptide substrate specificity, with inhibitors to a final concentration of 150 µM being added prior to incubation.

Table 3.3 Protease inhibitors used in this study

Inhibitor (150µM)	Target peptidase inhibition
Antipain Dihydrochloride	Trypsin
Chymostatin	Serine protease/ Chymotrypsin
Leupeptin	Serine proteases
TPCK	Chymotrypsin-like serine proteases
EDTA	Metallo-proteases
Pepstatin A	Aspartic proteases
E-64	Cysteine proteases

3.3 Results and Discussion

3.3.1 Cloning of the open reading frame encoding *CulsonLTRYP* into the pColdIII expression vector

In order to investigate the proteolytic properties of recombinant *C. sonorensis* late trypsin-like protease A, the nucleotide open reading frame had to be cloned into the expression vector, pColdIII. The purchased open reading frame nucleotide sequence of *C. sonorensis* late trypsin-like protease A, now labeled as *CulsonLTRYP*, was cloned into the pUC57 simple cloning vector between the restriction sites, NdeI and EcoRV, within the multiple cloning site of the vector. Upon receipt, the pUC57*CulsonLTRYP* DNA was rehydrated according to the method described by the manufacturer. In order to clone *CulsonLTRYP* into pColdIII, pUC57*CulsonLTRYP* and pColdIII were subjected to restriction digestion as described in Section 3.2.7. The digestion of pUC57*CulsonLTRYP* was successful, as two DNA fragments were clearly visible on an agarose gel (Figure 3.6 A, Lane 2). The digested pColdIII expression vector was visualised on an agarose gel shown in Figure 3.6 B, Lane 2. Following restriction digestion of pUC57*CulsonLTRYP*, two DNA bands were expected. The one band being the cloning vector, pUC57, with a size of around 2400 bp and the other *CulsonLTRYP* with a size of 878 bp. The expression vector pColdIII was expected to be around 4300 bp following restriction digestion.

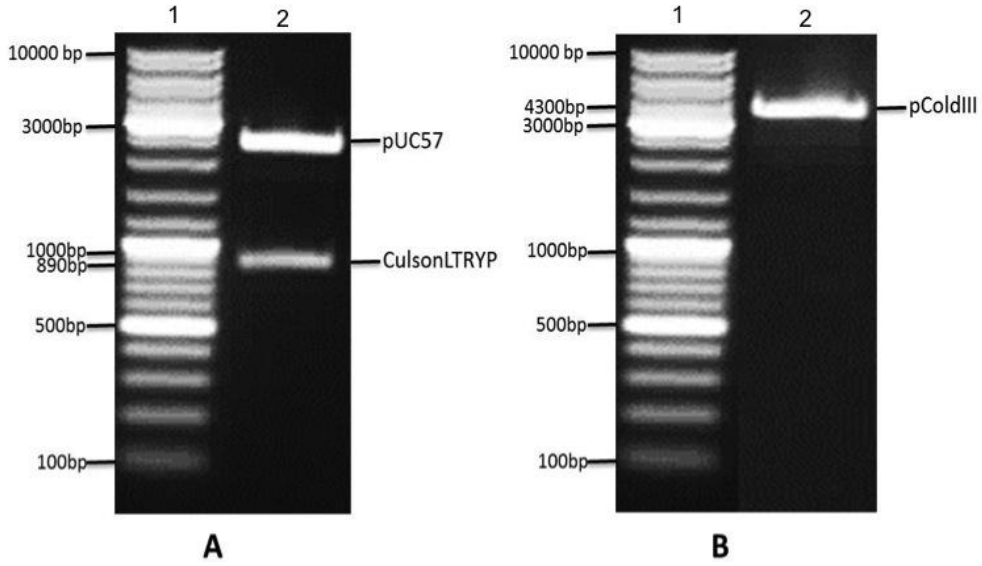

Figure 3.6 Agarose gel electophoretic analysis of the pUC57*CulsonLTRYP* (A) and pColdIII (B) vectors after digestion with NdeI and SalI restriction enzymes. Figure A: Lanes: 1) 8 µl O' GeneRuler DNA marker; 2) digested pUC57*CulsonLTRYP*. Figure B: Lanes: 1) 8 µl O' GeneRuler DNA marker; 2) digested pColdIII expression vector.

Both the pColdIII plasmid and insert were excised from the gels and DNA extracted as described in Section 3.2.9. The insert was ligated into the pColdIII vector and transformation yielded several colonies of transformants. In order to determine if the colonies contained the desired recombinant plasmids, colonies were screened by means of PCR using primers designed based on the insert sequence (Table 3.3). Colonies were considered positive if the amplicon visualised on an agarose gel had a size of around 890 bp. An example of such a screening can be seen in Figure 3.7. Restriction digestion was performed on isolated plasmid DNA of colonies considered to be positive from PCR screening. This was done in order to create complementary ends of the vector and the insert. The restriction digestion also facilitated the separation of the vector DNA from the *CulsonLTRYP* DNA on an agarose gel, which in return could be excised from the agarose gel.

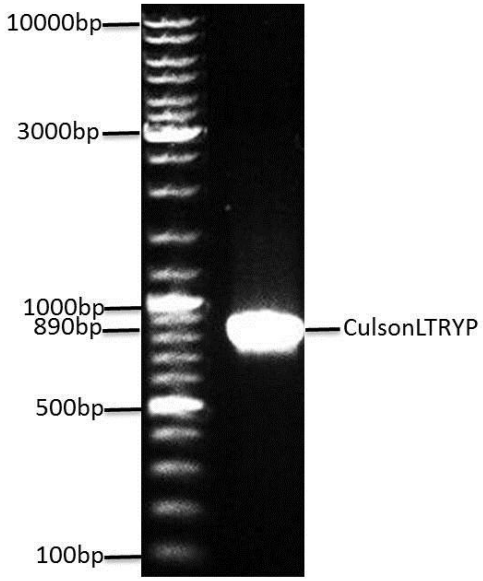

Figure 3.7 Agarose electrophoretic analysis of the PCR amplification of *CulsonLTRYP*. Lanes: 1) 8 µl O' GeneRuler DNA marker; 2) 890 bp CulsonLTRYP PCR amplicon.

At first glance, it seemed as if ligation was successful, as two DNA fragments of the expected sizes could be visualised on the agarose gel. As seen in Figure 3.8, the larger fragment of around 3400 bp was that of pColdIII and the smaller fragment, the insert, had a size of around 890 bp. However, the fact that there were noticeably more insert than vector, suggested that the cloning process might have had complications. One of these complications could be that the insert was double cloned into the pColdIII vector. To confirm this speculation, the isolated plasmids from positive colonies had to be sequenced.

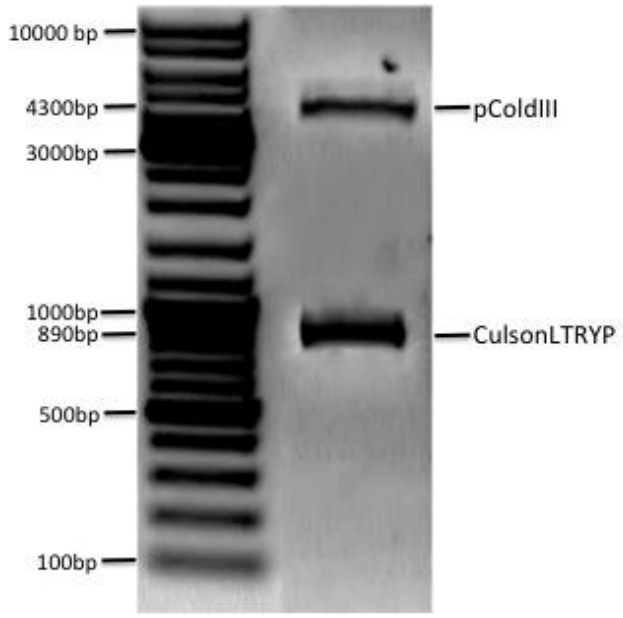

Figure 3.8 Agarose gel electrophoretic analysis of restriction enzyme screening of the expression vector containing *CulsonLTRYP* using NdeI and SalI. Restriction enzyme screening is performed to confirm colonies containing the desired recombinant plasmid. Lanes: 1) 8 µl O' GeneRuler DNA marker; 2) SalI and NdeI digested pColdIII*CulsonLTRYP*.

In the final part of screening, isolated plasmids from selected positive colonies were sequenced. This was to confirm that the *CulsonLTRYP* gene was cloned correctly without any sequence differences. Since no sequence abberations occured, the cloning of *CulsonLTRYP* into pColdIII expression was believed to be sucessfull. Therefore, the next part of this chapter, namely to express the recombinant *CulsonLTRYP* in bacterial cells was commenced.

3.3.2 Bacterial expression of pColdIII*CulsonLTRYP*

An initial expression experiment for recombinant *CulsonLTRYP* (r*CulsonLTRYP*) was performed using 0,5 mM IPTG and for an induction period of 12 hours. The control samples contained an empty pColdIII expression vector and were subjected to identical conditions as r*CulsonLTRYP.* Expression was analysed using SDS-PAGE.

As seen in Figure 3.9, Lanes 2-4, no expressed r*CulsonLTRYP* was visible on the SDS-PAGE gel. The total and soluble fractions of both the r*CulsonLTRYP* and control samples had similar expression profiles. The purified r*CulsonLTRYP*

sample also yielded no visible expressed r*CulsonLTRYP* indicating very little to no expression of the target protein. Furthermore, The only protein visible on the SDS-PAGE gel after purification was an unidentified protein with a size of around 40 kDa. Some natural bacterial host proteins have the ability to bind non-specifically to the matrix of purification columns, resulting in substandard purification.

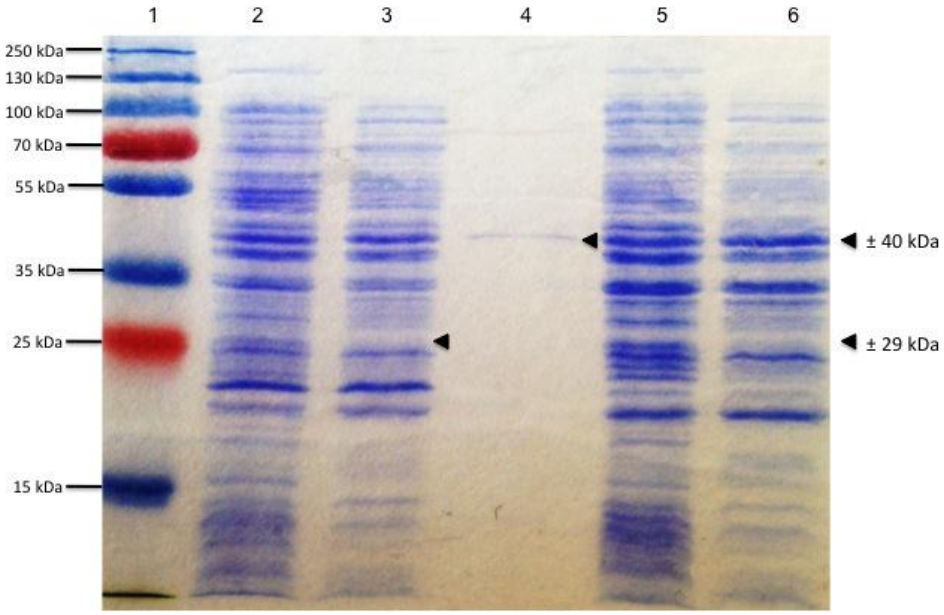

Figure 3.9 SDS-PAGE analysis of the expression of recombinant *CulsonLTRYP* at 0,5 mM IPTG and 12 h expression. Lanes: 1) 6,5 µl PageRuler protein marker; 2) pColdIII*CulsonLTRYP* total fraction; 3) pColdIII*CulsonLTRYP* soluble fraction; 4) pColdIII*CulsonLTRYP* partially purified soluble fraction; 5) pColdIII total fraction; 6) pColdIII soluble fraction.

It was reasoned that by altering the growth conditions, such as prolonging the expression period and increasing the concentration of the inducer, possible expression of r*CulsonLTRYP* could be expected. By increasing the induction time (16 hours) and the concentration of inducer (0,75 mM), more expressed proteins were visible on SDS-PAGE (Figure 3.10, Lane 3). The culture volume was also increased and the elution volume decreased which resulted in a higher concentration of expressed proteins being visible after partial purification. There were no obvious differences between the expression profiles of the r*CulsonLTRYP* sample (Lane 3) and the control samples (Lane 2). The proteins purified showed very low expression levels. Therefore, the final eluate of the purification had to be concentrated significantly using ultrafiltration (Lane 4). Unfortunately, by

concentrating this eluate, unwanted co-purified proteins such as the 40 kDa unidentified cellular protein (Lane 4) also became more noticeable when viewed on the SDS-PAGE gel. When comparing the induced *CulsonLTRYP* (Lanes 3-5) with the uninduced sample (Lane 2), it is visible that there was in fact expression of proteins in the sample.

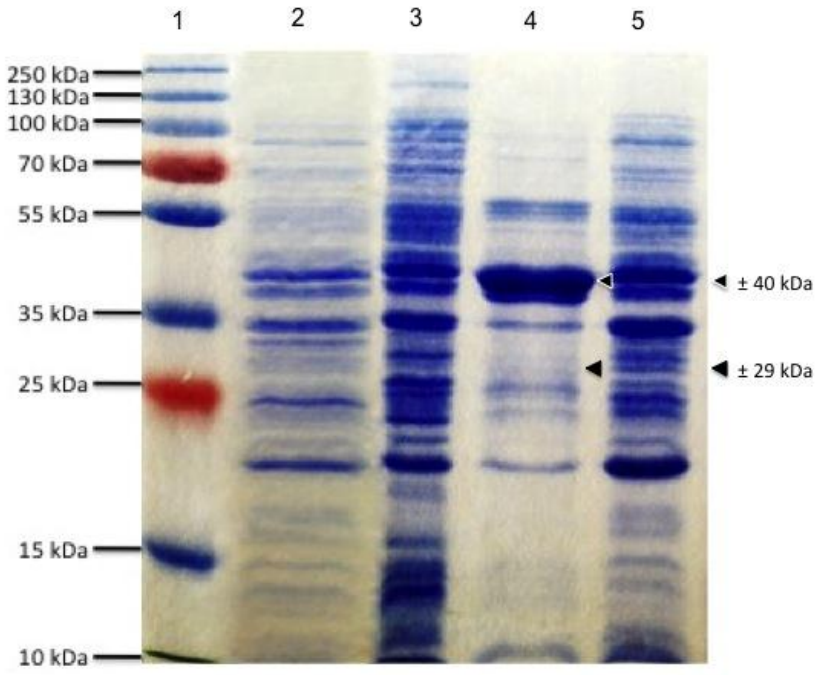

Figure 3.10 SDS-PAGE analysis of the expression of recombinant *CulsonLTRYP* at 0,75 mM IPTG and 16 h expression. Lanes: 1) 6,5 µl PageRuler protein marker; 2) Uninduced pColdIII*CulsonLTRYP*; 3) pColdIII*CulsonLTRYP* soluble fraction 4) pColdIII*CulsonLTRYP* partially purified and concentrated soluble fraction; 5) pColdIII soluble fraction.

Figure 3.11 show that by even further increasing the concentration of the inducer (1 mM) and expression time (24 h), higher levels of protein expression was visible on the SDS-PAGE gel. A very low yield of expressed r*CulsonLTRYP* was visible in the soluble fraction (Lane 3), which was absent in the total fraction (Lane 2) and control samples (Lane 5). By partially purifying the soluble fraction, considerable amounts of recombinant r*CulsonLTRYP* were lost. Even with concentration of the eluate, very little of the 29 kDa protein was visible on the SDS-PAGE gel (Lane 4).

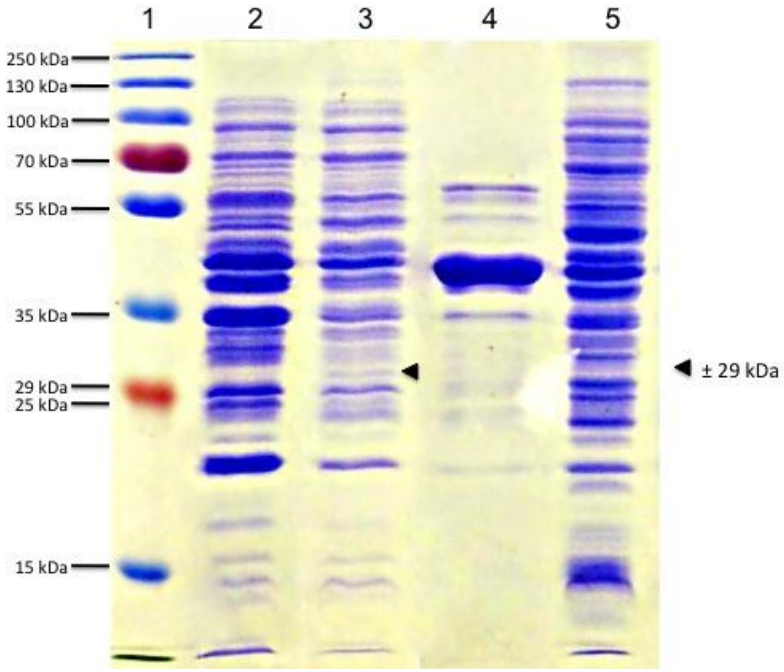

Figure 3.11 SDS-PAGE analysis of the expression of recombinant *CulsonLTRYP* at 1 mM IPTG and 24 h expression. Lanes: 1) 6,5 µl PageRuler protein marker; 2) pColdIII*CulsonLTRYP* total fraction; 3) pColdIII*CulsonLTRYP* soluble fraction; 4) pColdIII*CulsonLTRYP* partially purified and concentrated; 5) pColdIII soluble fraction.

Separation of the proteins with SDS-PAGE and staining with Coomassie blue was not sufficient for judging possible expression of the 29 kDa *CulsonLTRYP*. For this reason, a more sensitive analysis, namely silver staining, was used. A 29 kDa protein is clearly visible in Figure 3.12, Lane 2. Because of the high sensitivity of silver staining, several contaminating protein bands also became prominent on the SDS-PAGE gel.

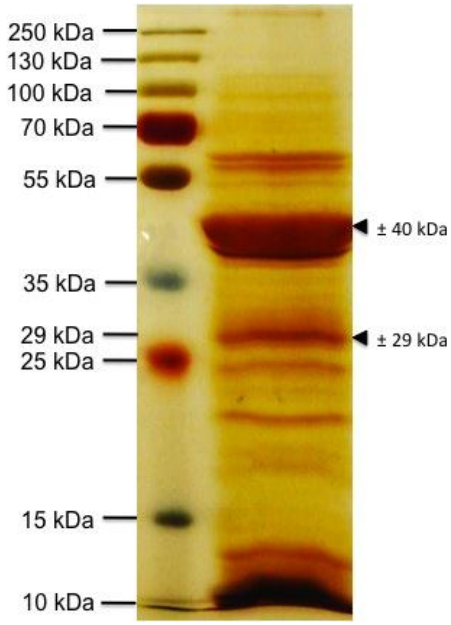

Figure 3.12 SDS-PAGE analysis of the partially purified and concentrated expressed recombinant *CulsonLTRYP* using silver stain visualisation. Lanes: 1) 6,5 µl PageRuler protein marker; 2) pColdIII*CulsonLTRYP*.

3.3.3 Determination of proteolytic activity of expressed recombinant *CulsonLTRYP*

The main aim of the research presented in this chapter was to determine if the late trypsin-like protease A of *C. sonorensis* could be expressed has an active protease. To be able to perform kinetic analysis, the preparation of enzyme to be investigated had to be at least partially purified (Palmer, 2001). The r*CulsonLTRYP* contained a C-terminal histidine tag and flexible serine-glycine in front of the histidine tag. This was to enable free movement of the histidine tag to bind to the matrix of the column resin and facilitate purification using immobilised metal chelate affinity chromatography (Loughran *et al*., 2006). Purification was performed as described in Section 3.2.15.

Proteolytic activity using Zymography (Gelatin-based substrate SDS-PAGE)

By incorporating gelatin into the resolving gel of an acrylamide gel (Section 2.2.4), proteases used the gelatin as a substrate and the proteolytic activity of the expressed r*CulsonLTRYP* was determined this way. Proteolytic activity of the

positive control sample, 0,005% (w/v) tryspin, was detected in Figure 3.13, Lane 6. Previous studies performed on the late trypsin-like serine protease of *C. sonorensis* established the size of the native protein to be around 29 kDa (Darpel *et al*., 2011). Trypsin is also a serine protease with a molecular weight of around 23 kDa (Figure 3.13, Lane 6). *CulsonLTRYP* containing a serine-glycine linker and a histidine tag were cloned into pColdIII and therefore, the predicted size of the active translated recombinant protein was to be around 29 kDa.

As seen in Figure 3.13, r*CulsonLTRYP* was expressed as an active protease when incubated at 37°C overnight. Distinct proteolytic activity could be seen at 25 kDa and 29 kDa of the total fraction (Lane 2), soluble fraction (Lane 3) and the purified sample (Lane 4). Purification and concentration of the eluate did not seem to have any effect on the activity of the expressed r*CulsonLTRYP* since equal levels of activity could be seen at the total fraction, soluble fraction and the purified sample. The control, which contained only an empty pColdIII plasmid, has shown no signs of proteolytic activity (Lane 5). This clearly indicates that activity originated from the expressed r*CulsonLTRYP*.

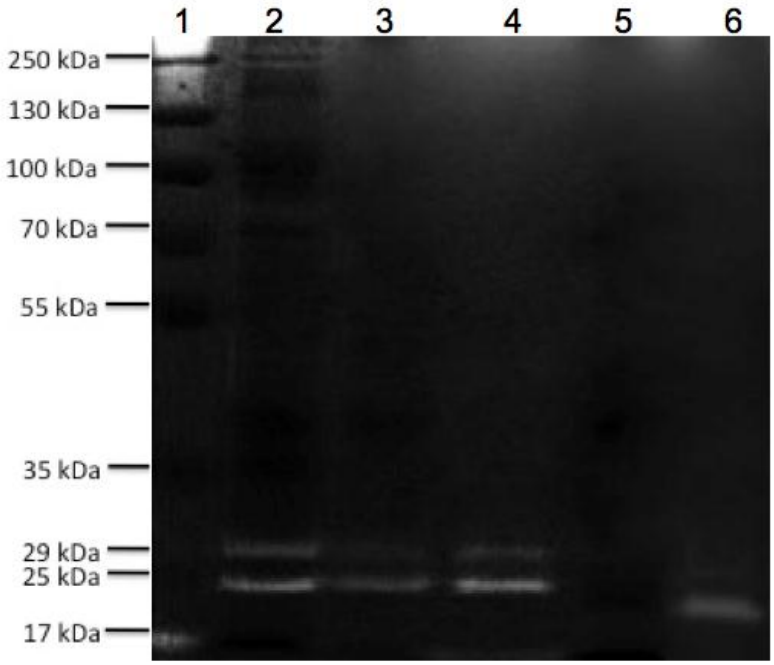

Figure 3.13 Gelatin based substrate SDS-PAGE analysis of the proteolytic activity of expressed recombinant *CulsonLTRYP*. Lanes: 1) 6,5 µl PageRuler protein marker; 2) pColdIII*CulsonLTRYP* total fraction; 3) pColdIII*CulsonLTRYP* soluble fraction; 4) pColdIII*CulsonLTRYP* purified and concentrated; 5) pColdIII soluble fraction; 6) 0.005% (w/v) trypsin activity control.

Since the partially purified sample had significant amounts of contaminants, the possibility that one of these contaminants being a protease existed. Trypsinogen is the inactive proenzyme of trypsin (Creighton *et al*., 1996). The presence of activity at two different sizes could possibly be explained by the partial cleavage of trypsinogen, resulting in two bands of activity being visible. From the results in Figure 3.13, it is clear that the expressed unidentified cellular protein had no influence on the activity of the expressed r*CulsonLRTYP* since no proteolytic activity was visible at the size of the unidentified cellular protein. Using zymography, active expression of r*CulsonLTRYP* was confirmed.

Fluorogenic peptide specificity

Different substrates, presenting basic residues at the P1 position and different residues in P2 and P3, were used to determine the substrate specificity of the expressed r*CulsonLTRYP*. This was decided upon as a result of previous studies characterising the substrate specificity for *C. sonorensis* late trypsin-like serine

protease A to be trypsin-like (Darpel *et al.*, 2011). Based on this, substrates presenting Arg residues at the P1 position were assayed in this study.

By incubating the expressed active r*CulsonLTRYP* in the presence of peptides containing a fluorescent AMC in P1, the proteolytic activity of the enzyme of interest could be determined. It is known that trypsin cleaves peptide chains predominantly at the carboxyl side of the amino acid Arginine (Arg), except when followed by Proline (Pro) (Rawlings & Barrett, 1994). It is shown in Figure 3.14 that r*CulsonLTRYP* accepted Arg at the P1 position of all the substrates. Arg was also the preferred residue at the P2 position as r*CulsonLTRYP* preferred Arg of BZ-Arg-AMC and CBZ-Ala-Arg-Arg-AMC over the Glycine (Gly) of Z-Pyr-Gly-Arg-AMC. As shown in Figure 4.14, r*CulsonLTRYP* had the lowest preference for substrates where Arg in P1 was followed by Pro in P2 such as Boc-Asp-(OBzl)-Pro-Arg-AMC and BOC-Val-Pro-Arg-AMC. The substrates BZ-Arg-AMC and CBZ-Ala-Arg-Arg-AMC are trypsin specific substrates demonstrating that the expressed r*CulsonLTRYP* probably has some trypsin-like characteristics.

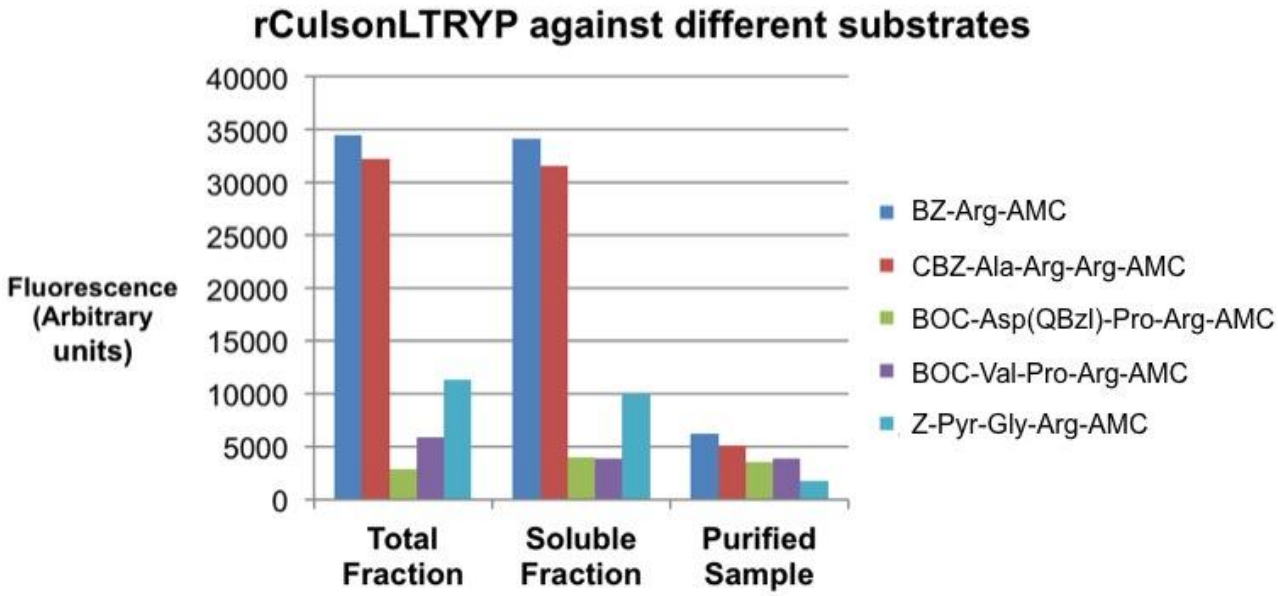

Figure 3.14 Enzyme assay of recombinant *CulsonLTRYP* against different substrates. The bars represent the intensity in which the expressed r*CulsonLTRYP* cleaves Arg at the P1 position resulting in the release of the fluorescent tag of the substrate. The graph represents the fluorescent units released at excitation of 360 nm and emission at 460 nm.

In the presence of a variety of substrates with different amino acids at the P1 position, the expressed r*CulsonLTRYP* favoured the cleavage of substrates with Arg at P1. Except when Arg was followd by Pro. As this is one of the major characteristics of trypsin (Rawlings & Barrett, 1994), it was believed that the partially purified protein had trypsin-like characteristics. However, to confirm this, the partially purified protein needed to be subjected to proteolytic inhibitors in order to better understand the activity profile of the expressed r*CulsonLTRYP*.

Inhibition profile of expressed r*CulsonLTRYP*

Protease inhibitors can be seen as compounds responsible for reducing the measured rate of enzyme-catalysed hydrolysis of a specific substrate. In the present study, it was important to conduct inhibitor profiling of the expressed r*CulsonLTRYP* to classify the expressed protease. Based on the strongest fluorescence profile obtained with the soluble fraction (Figure 3.14) and strong proteolytic activity shown in Figure 3.12, the soluble fraction of r*CulsonLTRYP* in conjunction with the BZ-Arg-AMC substrate was selected for the inhibitor profiling assays.

Figure 3.15 illustrates that antipain, an inhibitor that specifically targets the inhibition of trypsin had the strongest inhibition of r*CulsonLTRYP*. The serine protease inhibitors, chymostatin and leupeptin, also showed strong inhibition. TPCK is an inhibitor of chymotrypsin, with the exception that it does not inhibit trypsin (Salvesen & Nagase, 2001). Partial inhibition of r*CulsonLTRYP* was visible after incubation with TPCK and could possibly be explained by the excess impurities in the partially purified sample, which may have had chymotrypsin-like characteristics. r*CulsonLTRYP* is partially inhibited in the presence of the cysteine inhibitor, E-64. Even though E-64 has very weak abilities in inhibiting serine proteases, this inhibitor is still capable of inhibiting trypsin (Rawlings & Barrett, 1994; Katunuma & Kominami, 1995). Pepstatin A is a highly selective inhibitor of aspartyl proteases and is known for its inability to inhibit serine proteases (Umezawa, 1972). A partially weak inhibition of r*CulsonLTRYP* was visualised in the presence of pepstatin A and could possibly be explained by the presence of proteolytic impurities in the sample such as cathepsin D, renin, pepsin or bovine

chymosin (Marciniszyn, 1976). No inhibition was achieved when incubating r*CulsonLTRYP* with metallo-protease inhibitor, EDTA.

As the name suggests, C. *sonorensis* late trypsin-like serine protease A, was classified as a trypsin-like serine protease (Darpel *et al*., 2011). Therefore, protease inhibitors, specifically targeting trypsin and serine proteases were used for the possible inhibition of r*CulsonLTRYP* proteolytic activity. As shown in Figure 3.15, the trypsin specific inhibitor, antipain, had the strongest inhibition profile of all the inhibitors. Furthermore, the serine inhibitors, chymostatin and leupeptin also strongly inhibited *rCulsonLTRYP*, but less than antipain. Also, the salivary protease identified by Darpel *et al* (2011) was classified as a trypsin-like serine protease. Therefore, based on the results presented in this chapter, it was a good indication that the expressed r*CulsonLTRYP* was also a trypsin-like serine protease.

The primary function of TPCK is to inhibit chymotrypsin, even though it has been shown to have weak abilities in inhibiting trypsin as well (Carpenter, 1967). This could explain the relatively weak inhibition profile of r*CulsonLTRYP* seen in Figure 3.15, when compared to the stronger trypsin inhibitors. E-64 is responsible for the inhibition of cysteine proteases. This inhibitor is incapable of inhibiting serine proteases with the exception of trypsin (Barret *et al.,* 1982). This was seen by the partial inhibition of r*CulsonLTRYP* by E-64. Since the expressed r*CulsonLTRYP* sample was only partially purified, it was likely that other proteases were also present in the sample. Some of the proteases might have been aspartyl proteases, since a form of inhibition of r*CulsonLTRYP* was detected in the presence of pepstatin A. No inhibition was present when r*CulsonLTRYP* was incubated with EDTA. This indicates that the expressed r*CulsonLTRYP* is not a metallo-protease. Based on these findings, it seems that the recombinant protein expressed has trypsin-like characteristics which correlates to the characteristics of the to the trypsin-like serine protease reported by Darpel *et al* (2011).

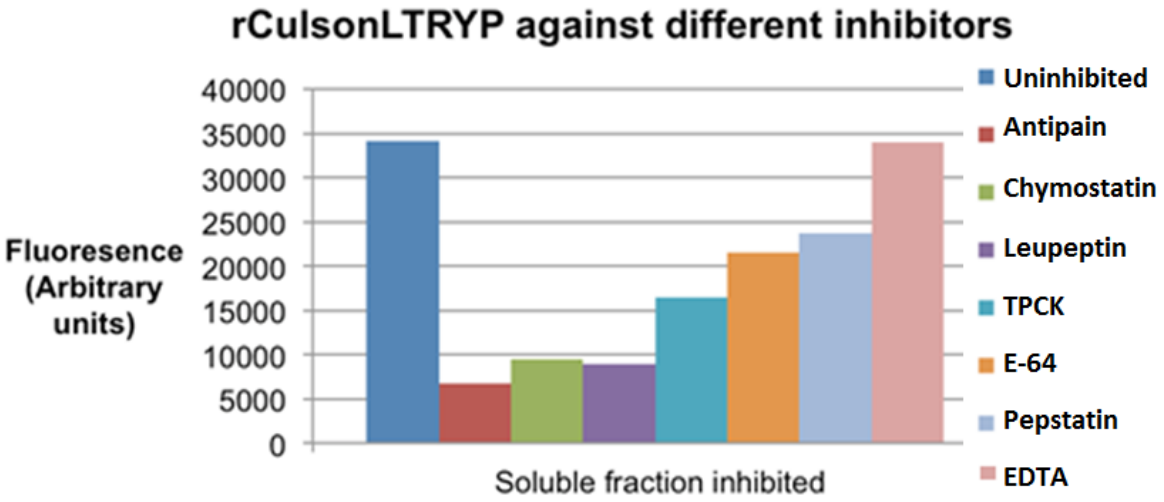

Figure 3.15 Inhibitor profiles of expressed r*CulsonLTRYP* in the presence of different protease inhibitors. BZ-Arg-AMC was used as substrate for r*CulsonLTRYP* and the legend indicates the different inhibitors used. The bars represent the fluorescent units released at excitation of 360 nm and emission of 460 nm.

3.4 Summary

This part of the study had three main objectives. The first objective was to clone *CulsonLTRYP* protein containing a histidine tag into the cold expression vector, pColdIII. Secondly, to actively express the recombinant *CulsonLTRYP* in bacteria and finally to verify the proteolytic activity of the expressed enzyme using different enzymatic assays.

The nucleic acid sequence for the 29 kDa late trypsin-like serine protease A from C. *sonorensis* was obtained from GenBank (Accession number: AY603563.1). A histidine tag was added to the C-terminus using a serine-glycine linker. The linker provided flexibility to the histidine tag allowing easier binding of the histidine tag to the matrix of the metal chelate affinity purification column (Macherey-Nagel, Purification of His-tag proteins: User manual). The plasmid and insert were digested using restriction enzymes and *CulsonLTRYP* was ligated between the restriction sites SalI and NdeI of pColdIII. An initial expression period of 12 hours and inducer concentration of 0,5 mM IPTG yielded no visible soluble protein. The proteins expressed had a similar profile to the expressed empty pColdIII vector

that served as the control. The soluble fraction was purified using metal chelate affinity chromatography and only a single expressed protein of around 50 kDa was visible. It was believed that some cellular proteins were bound non-specifically to the matrix column resulting in the co-elution of unwanted proteins with the target protein. The increase in expression time and a higher inducer concentration produced a visibly higher amount of proteins. Concentration of the purified eluate resulted in a higher concentration of the unwanted co-purified proteins being visualised by SDS-PAGE. An increase in inducer concentration and expression time seemed to contribute to the expression of more proteins. Doubling the initial inducer concentration resulted in the expression of a soluble protein band with the predicted 29 kDa size of *CulsonLTRYP*. Purification diluted the expressed protein of interest leading to difficulty in visualising the expressed protein. By using the highly sensitive protein visualisation technique, silver staining, the expressed *CulsonLTRYP* could clearly be visualised.

The final objective of this part of the study was to determine if the expressed r*CulsonLTRYP* was expressed as an active enzyme. The proteolytic activity of r*CulsonLTRYP* was verified by means of zymography. As previously stated *CulsonLTRYP* is characterised as a trypsin-like serine protease. Proteolytic activity of the expressed r*CulsonLTRYP* was observed at two positions of the zymogram, namely at 25 kDa and 29 kDa. Activity at these sizes correlates with the expected size of r*CulsonLTRYP* (Darpel *et al*., 2011). Based on these findings, a suitable substrate needed to be identified to be used in further fluorescence analyses. The substrates for r*CulsonLTRYP* were selected based on this principle. Substrates having an Arg at the P1 position showed high levels of fluorescence, in contrast to the substrates where Arg was followed by Pro (Rawlings & Barrett, 1994). This finding was a good indication that the expressed r*CulsonLTRYP* has trypsin-like activity. BZ-Arg-AMC had the highest level of fluorescence and was used as substrate in subsequent assays. To finally confirm that the expressed r*CulsonLTRYP* was trypsin-like, r*CulsonLTRYP* was incubated in the presence of a variety of protease inhibitors. These inhibitors ranged from trypsin-, serine-, cysteine- and metallo-protease inhibitors. Inhibition of r*CulsonLTRYP* was achieved by all of the inhibitors capable of inhibiting serine proteases. The trypsin inhibitor, antipain had the strongest inhibition. The unexpected partial inhibition of

r*CulsonLTRYP* by pepstatin A could be explained by the presence of impurities containing aspartyl proteases in the analysed sample (Tang & Wong, 1987). No inhibition was achieved by EDTA confirming that r*CulsonLTRYP* is not a metallo-protease.

Some natural bacterial host proteins have the ability to bind non-specifically to the matrix of the column, resulting in the co-elution of these bacterial proteins causing sub-optimal purification. This was most likely the case with the higher molecular mass unknown cellular protein seen in Figures 3.9 - 3.12. The conditions for expression were ideal for this unknown cellular protein, which resulted in the high level of expression visible. This protein did not show any proteolytic activity on the gelatin SDS-PAGE. Therefore, it was believed that this unknown cellular protein would not influence further analysis with the partial pure preparation and additional steps in purification were not followed. Very low levels of expression were achieved which could be explained by the contaminating host proteins that bound better to the resin than the small amounts of r*CulsonLTRYP.* Since the low amounts of polyhistidine-tagged *CulsonLTRYP* protein were not able to replace the majority of the contaminating proteins, alternative measures need to be followed for optimal expression. Some of these expression steps should include the increase of the expression level and avoid exceeding the recommended lysis volumes. Purification could also be improved by including additional wash steps. This should rid the sample of any unwanted contaminating proteins (Saluta & Bell, 1998). Furthermore, small amounts of imidazole could also be added which would help with the binding of the histidine protein to the matrix. Finally, the addition of a small amount of β-mercaptoethanol would lower the disulfide bonds that may be linking r*CulsonLTRYP* with the contaminants, resulting in fewer contaminants being co-eluted with r*CulsonLTRYP* (Saluta & Bell, 1998). These modifications may aid in future investigations of the various properties of this enzyme. The development of an expression and purification system for biologically active recombinant r*CulsonLTRYP* will also contribute to the gaining of knowledge for this enzyme. This knowledge could be used to aid in the rational design of therapies for AHSV.

Based on the work done in this chapter, it can be concluded that the recombinantly expressed *CulsonLTRYP* is biochemically active and has trypsin-like activity. The expressed r*CulsonLTRYP* has also shown similar characteristics in activity when compared with the proteins of *C. imicola* identified in Chapter 2. The work done in this chapter will contribute to the next goal of this study in which the effect of recombinant protease digestion on AHSV4 VP2 will be investigated.

CHAPTER 4

Digestion of AHSV4 using a recombinant 29 kDa *C. sonorensis* protease

4.1 Introduction

Culicoides salivary proteases (Darpel *et al.*, 2011) have the ability to cleave the structural proteins of AHSV (Marchi *et al*., 1995). This cleavage could lead to a possible increase in infectivity for the insect vector (Marchi *et al*., 1995). Darpel *et al* (2011) identified a trypsin-like serine protease in the saliva of the BTV vector, *C. sonorensis,* resulting in structural changes of the virus. Since this vector can act as host for both BTV and AHSV (Boorman *et al*., 1975; Wellby *et al*., 1996), and BTV and AHSV are structurally similar (Howell, 1962), the effects of a recombinant *C. sonorensis* protease digestion on African horsesickness virus serotype 4 (AHSV4) needed to be investigated.

Previous studies performed on BTV and AHSV have demonstrated proteolytic cleavage of the outer capsid protein, VP2, in the presence of trypsin or chymotrypsin (Van Dijk & Huismans, 1982; Darpel *et al*., 2011). This cleavage results in the formation of ISVPs (Van Dijk & Huismans, 1982; Marchi *et al*., 1995). Previous studies have reported that the selective cleavage of the outer capsid protein VP2 by the action of trypsin-like serine proteases would result in the formation of these ISVPs, which in return circulate the *Culicoides* species that vector AHSV (Marchi *et al*., 1995). It was also reported that the enzymatic cleavage of AHSV9 particles enhanced the infectivity of the virus in insects by up to 1000-fold (Marchi *et al*., 1995). It is therefore important to determine what effect recombinantly expressed trypsin-like serine proteases would have on the cleavage pattern of AHSV VP2.

The two main aims of this part of the project were to firstly cultivate and purify AHSV4 and secondly, to subject the purified virus to the partially purified

expressed r*CulsonLTRYP* protease (Chapter 3) and analyse the cleavage patterns of VP2 using SDS-PAGE.

4.2 Materials and Methods

4.2.1 Culturing of the BHK-21 cell line

The mammalian cell line used in this part of the study was BHK-21 (Clone 13) (Figure 4.1). This cell line was originally derived from the single cell isolation from kidneys of 1-day-old baby Syrian hamsters (*Mesocricetus auratus*) (Stoker & Macpherson, 1962; Erasmus, 1964). BHK-21 cells are used extensively for AHSV propagation (Mirchamsy & Rapp, 1968) and therefore, this specific cell line was used in this study.

The cells to be used in this study were generously provided by Prof. A.C. Potgieter (Deltamune, Pretoria, South Africa). Upon receipt, the two 75cm^3 flasks contained cells at passage number 38 and were around 70% confluent. The cells were incubated in a humidified incubator (37°C, 5% CO_2) until they reached 90% confluency.

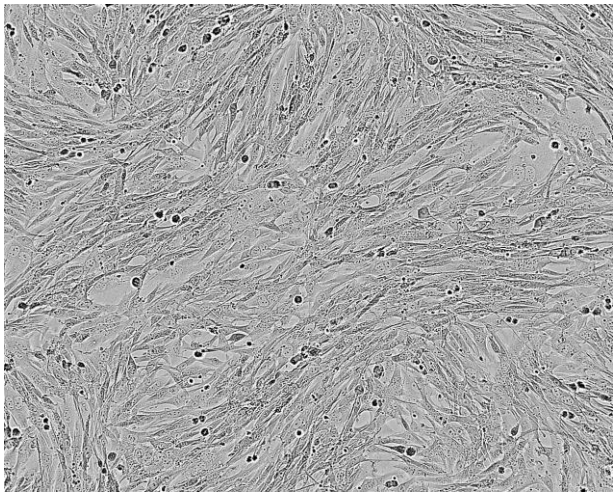

Figure 4.1 BHK-21 cells at around 85% confluency (http://hpacultures.org.uk: Public health England). This cell line was derived from the kidneys of baby hamsters and is frequently used in virus propagation studies.

The entire spectrum of cell culturing procedures was carried out in a laminar flow hood using sterile solutions and equipment. Before start of passaging, fresh culture medium was prepared. The basal medium used was Dulbecco's modified eagle medium (DMEM), a medium that allows the culture of mammalian cells with reduced foetal bovine serum (FBS) supplementation (Morgan *et al.*, 1950). Instead of the usual 10% FBS, 5% was used, since higher percentages caused the cells to grow too fast (Rahman *et al.*, 2006). L-glutamine was used as a cell supplement for protein synthesis and glucose production as well as the formation of amino sugars and glutathione. Cell growth and viability was increased with the supplementation of non-essential amino acids (NEAA) (Rahman *et al.*, 2006). The antibiotics penicillin and streptomycin were used to prevent bacterial contamination of the cell culture. These antibiotics have a highly effective combined action against Gram-positive and Gram-negative bacteria (Schatz *et al*., 1944). By supplementing the antibiotics with Amphotericin B, fungal contamination of the cell cultures was prevented. Unless otherwise stated, the culture medium consisted of DMEM supplemented with 0,5 mM NEAA, 0,5 mM L-glutamine, 60 mg/ml penicillin, 60 mg/ml streptomycin, 150 µg/ml amphoterocin B and 5% FBS.

When the cells reached 90% confluency, the medium was decanted and the cells washed once with 10 ml PBS. This was to ensure the complete removal of the serum, which contained trypsin inhibitors (Jacobsson, 1955). A 1x trypsin solution was added to the monolayer and incubated for about 2 minutes at room temperature. Following incubation, 90% of the trypsin solution was decanted and the status of cell detachment monitored under an inverted microscope. The cells were incubated until most of the cells were detached. Trypsination was stopped through the addition of culture medium to the cells. BHK-21 cells are notorious for clumping and clumps that formed were broken up by briefly pipetting the solution. Again, the cells were monitored under the microscope and cell viability was confirmed to be over 90%.

The cells were generally maintained in 75 cm^2 flasks using a splitting ratio of 1:10. This was achieved by transferring 1 ml of the cell suspension to a new 75 cm^2 flask containing 10 ml culture medium. To expand the cells, 3 ml of the cell suspension was added to 45 ml culture medium in a 175 cm^2 flask. The cells were

incubated in a humidified incubator (37°C, 5% CO_2) until 90% confluency was reached, after which the next passage was performed.

Freeze preservation of mammalian cells is often used to minimise genetic change in continuous cell lines, to avoid ageing and transformation or loss of cells through contamination (Freshney, 1987). Before freezing, the cells were checked for any signs of contamination and the cryopreservation medium [DMEM containing 2 mM L-Glutamine, 10% (v/v) FBS, 10% (v/v) DMSO] was prepared (Stoker & Macpherson, 1962). The cryopreservative, dimethyl sulfoxide (DMSO), and L-Glutamine was added to the cryopreservation media to protect the cells from the freeze-thaw process (Freshney, 1987). When the monolayer cells reached 90% confluency, the cells were trypsinated after which DMEM was added to the cells. The cells were counted under an inverted microscope using a haemacytometer. The cells were collected by centrifugation (2 000 *g*, 5 minutes, 15°C) and resuspended in cryopreservation medium at a density of 3 x 10^6 cells/ml. Aliquots of 1 ml were prepared and the cells were stored at -80°C.

4.2.2 Propagation of AHSV4

The virus to be used in this study was generously provided by Prof. A.C. Potgieter (Deltamune, Pretoria, South Africa). The strain used was from the AHSV4 reference strain which was originally adapted in Vero cells. The virus number was HS 32/62 1S- 2 Vero. This means that this specific strain was the 32_{nd} isolate for AHSV during the year of 1962 at the ARC-OVI. This isolated strain was passaged once in baby mice and twice in Vero- E6 cells by the Deltamune team (Personal communication with Prof. A.C. Potgieter).

A 75 cm^3 flask containing 90% confluent monolayer BHK-21 cells was propagated as described in Section 4.2.1. The medium was replaced with a fresh volume of culture medium including 1 ml of the first passage AHSV4 inoculum. The culture was incubated in a humidified incubator (37°C, 5% CO_2) and cell growth monitored under an inverted microscope until full cytopathic effect (CPE) of the cells was reached. The second-passage virus inoculum was transferred to a sterile tube and stored at 4°C until use.

4.2.3 Virus quantification by $TCID_{50}$ titration

The quantification of viruses is practiced widely in both basic and applied virology. The concentration of an infectious virus sample can be measured through several methods. One of the most common approaches to quantify the concentration of virus is by means of the tissue culture infectious dose at 50% assay ($TCID_{50}$) (Reed & Muench, 1938). In short, cultures of confluent cells susceptible to infection are treated with a series of viral dilutions and scored according to whether or not cell death has occurred. By using the formula published by Reed and Muench (1938), the titer of the virus stock can then be determined and is expressed as the $TCID_{50}$/ml. The $TCID_{50}$/ml value represents the concentration of virus required to cause cell death or pathological transformations in 50% of the inoculated cell cultures.

The cells that were used in the titration were harvested from a 75 cm^3 flask containing a monolayer of 90% confluent cells. The second-passage virus inoculum was serially diluted with PBS up to 1:10 and the tubes were labeled from 10^{-1} to 10^{-10}. A 96-well plate was labeled by drawing grid lines to delineate quadruplicates and each grid was numbered to correspond to the virus dilution. A volume of 80 µl of the harvested cells was added to each of the wells and incubated in a humidified incubator (37°C, 5% CO_2, 2 hours). This was to ensure attachment of the cells to the plate before infection. Starting from the most dilute sample, 20 µl virus was added to each of the quadruplicate corresponding wells. The titration mixture was incubated in a humidified incubator (37°C, 5% CO_2, 4 days) and the CPE was monitored using an inverted microscope.

The purpose of staining biological samples is to enhance the visible contrast between the structures in the cell under the microscopic image (Fox, 1977). Following incubation, the medium was decanted from the 96-well plate and the cells were gently rinsed under running water. Live cells unaffected by the virus remained attached to the plates, whereas cells that died from virus infection became detached. This resulted in the dead cells being washed away during rinsing. The colouring agent used to visualise the cells was trypan blue. Trypan blue is only capable of traversing the membrane of cells that are dead (Strober,

2001) and therefore, the live cells were fixed with 100% (v/v) ethanol for 5 minutes at room temperature. The ethanol was decanted. The remaining cells that died from fixation were coloured with a 0.1% (w/v) trypan blue solution. Following staining, the stain was decanted and the cells were gently washed with running water. The cells unaffected by the virus was visualised as distinct blue cells and colourless wells were observed where the virus had a cytopathic effect on the cells. The number of positive and negative wells was recorded on a light table and the $TCID_{50}$/ml was calculated (Reed & Muench, 1938).

4.2.4 Purification of AHSV4

AHSV4 was purified according to a method described for the purification of BTV (Huismans *et al.*, 1987). Confluent monolayers BHK cells were grown as described in Section 4.2.1 and inoculated with the second passage AHSV4 inoculum (Section 4.2.2). The cells were incubated for around 22 hours or up to just before CPE started.

Prior to the harvest of the virus-infected cells, the glass beads to be used for harvest were sterilised. This was done by boiling the glass beads (3 mm diameter) in 0,1 M HCl for 30 minutes, rinsed in several changes of distilled water and autoclaved (121°C, 1.1 Bar, 20 minutes). The inoculated cells were harvested using the glass beads and collected in volumes of 250 ml in conical flasks by centrifugation (2000 *g*, 20 minutes, 4°C). All low speed centrifugation steps were performed in a tabletop centrifuge at 4°C. After removal of the supernatant, the cells were resuspended in 10 ml lysis buffer [2 mM Tris-HCl buffer, 0,5% Triton X-100, pH 8.8] and incubated on ice for 10 minutes. The cells were then disrupted with 10 strokes of a tight fitting dounce glass homogeniser. The cell nuclei and debris were collected by centrifugation (5000 *g*, 10 minutes, 4°C) and washed twice in 10 ml lysis buffer. The cytoplasmic extracts were combined after which the virions were recovered by ultracentrifugation (131 000 *g*, 90 minutes, 4°C) through a 5 ml sucrose cushion [2 mM Tris-HCl, 40% (w/v) sucrose, pH 8.8]. The supernatant and sucrose cushion were gently aspirated after which the partially purified virus pellet was resuspended in 1 ml Low Tris Buffer (LTB) [2 mM Tris-HCl buffer, pH 8.8]. The boyant density of Orbiviruses is around 1,36 g/ml and can

therefore be relatively easy purified by centrifugation through a sucrose gradient (Roy, 2011). A 4-40% sucrose gradient was prepared on a Biocomp gradient station gradient mixer (Biocomp, Canada) using 4% and 40% (w/v) sucrose solutions in LTB. The resuspended virus pellet was ultracentrifuged (131 000 *g*, 40 minutes, 4°C) through the sucrose gradient. The visible virus band that formed was collected with an 18 G syringe and diluted with approximately five volumes LTB. The diluted virus was ultracentrifuged (131 000 *g*, 90 minutes, 4°C) to concentrate the virus. The supernatant was aspirated and the virus pellet resuspended in 500 µl LTB containing sucrose [2% (w/v) Sucrose]. The partially purified virus with a final volume of 500 µl was visualised using SDS-PAGE according to the method described in Section 2.2.3. The virus preparations were stored at 4°C until further use.

Caesium chloride (CsCl) is one of many materials that can be used for the purification of viruses (Szybalski, 1960). It is regularly used in a technique referred to as density-gradient centrifugation. In short, this technique is based on the buoyant density separation of particles. The virus suspension is added to the self-forming gradient mixture and ultracentrifuged whereupon the viral particles will reach equilibrium according to their buoyant density. CsCl is a water-soluble substance that forms a gradient at ultra-high speeds enabling the migration of nucleic acids up to the point in which they reach neutral buoyancy (Carr & Griffiths, 1987).

A 2 M CsCl stock was prepared with LTB after which the refractive index (RI) was determined using a refractometer. Virions in the family *Reoviridae* have buoyant densities in CsCl of 1,36 - 1,39 g/cm^3 (Roy, 2011). The partially purified virus sample was mixed with the CsCl mixture and the RI adjusted with the 2 M CsCl stock solution until the RI was 1,3651 g/cm^3. The sample was split into two and ultracentrifuged (129 000 *g*, 18 hours, 18°C). Bands that formed at equal densities in the split samples were collected under slit light and pooled. Residual CsCl was removed by diluting the collected virus with 5 ml LTB. The virus was pelleted by ultracentrifugation (131 000 *g*, 90 minutes, 4°C). The supernatant was aspirated and the virus pellet resuspended in 250 µl LTB containing sucrose [2% (w/v)

Sucrose]. The virus preparation was visualised using SDS-PAGE (Section 2.2.3) and virus sample stored at 4°C until further use.

4.2.5 Proteolytic digestion of AHSV4 with r*CulsonLTRYP*

To investigate the possible cleavage of AHSV4 VP2 through incubation with the expressed recombinant trypsin-like serine protease, r*CulsonLTRYP*, 20 µl of the partially purified AHSV4 preparation was added to 10 µl r*CulsonLTRYP*. The reaction mixtures were incubated in a circulating water bath set at 37°C at time intervals of 1 h, 3 h and 6 h respectively. The digested virus-recombinant protease preparations were stored at 4°C until further analysis. Proteins in the reaction mixtures were denatured with β-mercaptoethanol and boiled for 3 minutes. The protein fragments were analysed on SDS-PAGE gel that was silver stained according to the method described in Section 2.2.3

4.3 Results and Discussion

4.3.1 Preparation and titration of AHSV4 using $TCID_{50}$/ml virus quantification

The infectious titer of the second passage AHSV4 inoculum was determined by infecting in quadruplicate a 96 well plate containing BHK-21 cells with a series of virus dilutions. The $TCID_{50}$/ml was calculated using the formula by Reed and Muench (Reed & Muench, 1938). Following titer determination, virus was cultivated in order to be purified for subsequent analyses.

Based on the colouring of the BHK-21 cells, the wells could easily be scored as positive or negative (Table 4.1). The wells that stained blue was scored as negative for virus infection since the cells in the wells were alive and did not detach as a result of viral infection. The unstained wells were indicative of virus infection since the cells detached from the plate and could subsequently not be stained. These wells were scored as positive for infection (Figure 4.2). Virus sample diluted from 10^{-1} to 10^{-4} was positive, as full CPE was reached. The 10^{-5} dilution had two positive wells and at the 10^{-6} dilution six out of the eight wells

reached full CPE. An equal number of viable cells and cells that reached full CPE was visible at the 10^{-7} virus dilution. A noticeable decline in the ability of the virus to cause CPE in the cells was observed from dilution 10^{-8} onwards. At the 10^{-8} virus dilution five wells were negative and from virus dilutions 10^{-9} to 10^{-10} no signs of CPE were visible.

Table 4.1 Percentage infection and infection ratio from noted positive and negative wells

Virus dilution	Infections per number of wells inoculated	Observed values		Cumulative values		Infection ratio	% Infection
		Positive	Negative	Positive	Negative		
10^{-1}	8(8)	8	0	47	0	47(47)	100
10^{-2}	8(8)	8	0	39	0	39(39)	100
10^{-3}	8(8)	8	0	31	3	31(34)	91
10^{-4}	8(8)	8	0	23	7	23(30)	77
10^{-5}	2(8)	2	6	15	13	15(28)	54
10^{-6}	6(8)	6	2	13	15	15(28)	54
10^{-7}	4(8)	4	4	7	23	7(30)	23
10^{-8}	3(8)	3	5	3	31	3(34)	9
10^{-9}	0(8)	0	8	0	39	0(39)	0
10^{-10}	0(8)	0	8	0	47	0(47)	0

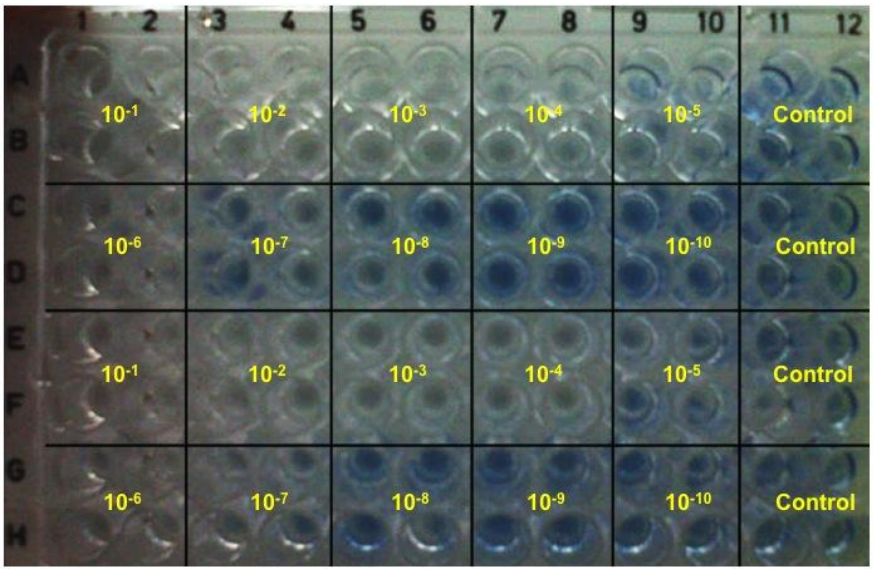

Figure 4.2 Titration plate used in $TCID_{50}$/ml titration assay indicating positive and negative wells after a 4-day incubation period. Rows A, B, E and F numbered 1 to 10 represents dilutions 10^{-1} to 10^{-5}. Rows C, D, G and H numbered 1 to 10 represent dilutions 10^{-6} to 10^{-10}. Rows A to H numbered 11 to 12 represents the control samples. The control sample did not contain any virus and the virus was replaced with PBS. Since the wells that stained blue contained live cells, these wells were noted as negative for infection. The cells in the unstained wells were dead as a result of viral infection and noted as positive for infection.

By tabulating the positive and negative results from Figure 4.2, the cumulative values for each dilution was determined. This enabled the calculation of the infection ratio and the percentage infection for each of the virus dilutions. Using the values in Table 4.1 in conjunction with the equation of Reed and Muench (1938) (Equation 4.1), the $TCID_{50}$/ml endpoint was calculated.

Equation 4.1 End point dilution by Reed and Muench (1938)

$$X = (Log_{10}\, Dilution\ factor)\frac{(\%\ \text{Infection at next dilution above } 50\%)\ -\ 50\%}{(\%\ \text{Infection at next dilution above } 50\%) -\ (\%\ \text{Infection at dilution below } 50\%)}$$

The proportional distance between the dilutions spanning 50% for the data shown in Table 4.1 gives

$$x\ =\ (1)\frac{(54-50)}{(54-23)}$$

Thus,

$$x = 0.13$$

This proportional factor was added to the next dilution above 50% ($10^{-(6 + x)}$), which would yield an end point dilution of $10^{-6.13}$ for every $TCID_{50}$ per 20 µl virus sample.

The reciprocal of this value was used,

$$Virus\ concentration = \frac{1}{10^{-6.13}}$$
$$= 1348962.883 / 20\ \mu l$$

along with the adjustment of the sample volume to 1 ml

$$Virus\ concentration = 1348962.883\ TCID_{50} / 0{,}02\ ml$$

to give the titer of the virus stock.

$$Virus\ concentration = 6.7 \times 10^7\ TCID_{50}/ml$$

The infectious titer of the virus was based on a series of dilutions used to calculate the $TCID_{50}$/ml through the use of the Reed and Muench equation (Reed & Muench, 1938). The viral concentration of the second generation AHSV4 inoculum resulted in a noticeable cytopathic effect on the cells after four days. As the virus was diluted over a large range, the infection capability of the virus diminished in some of the replicates, ultimately to the point in which the virus had no cytopathic effect on the cells in the set period of time. A dilution of 10^{-5} was needed to infect 50% of the cells over the given period of time, and through the use of a mathematic equation, the infection titer of the virus was calculated to be 6.7 x 10^7 $TCID_{50}$/ml virus.

By determining the infection titer, the volume of virus inoculum needed to infect cells for virus purification was established. The second passage virus preparation had a final volume of 15 ml and therefore the total infection titer of the second passage virus culture was 1.0 x 10^9 $TCID_{50}$/15ml. The $TCID_{50}$/ml titer value obtained was relatively low indicating that a high volume of inoculum was needed to infect cells. Therefore, to be able to study the effects of protease digestion on

the virus particle, more virus had to be cultivated using the infection titer value obtained by the $TCID_{50}$/ml value. This enabled the production of a larger virus work volume for subsequent protease-virus analyses.

4.3.2 Purification of AHSV4

Monolayers of confluent BHK-21 cells were infected using second passage AHSV4 with an infection titer of 6.7 x 10^7 $TCID_{50}$/ml (Section 4.3.1). At the first sign of CPE, the virus-infected cells were harvested, released from the cells into the cytoplasm, the virions were then recovered by ultracentrifugation and resuspended in LTB for SDS-PAGE analysis (Figure 4.4).

The expected sizes of the AHSV4 structural proteins were observed at 111 kDa (VP2), 103 kDa (VP3), 59 kDa (VP5) and 38 kDa (VP7) on the SDS-PAGE gel (Figure 4.3). The observation of protein bands consistent with the published sizes of structural proteins of AHSV4 in SDS-PAGE, confirmed that AHSV4 was cultivated in the BHK-21 (Manole *et al*., 2012). However, virus purification by means of sucrose gradient ultracentrifugation was sub-standard as there were a significant amount of impurities present in the virus preparation. Even though protein bands at the expected sizes of AHSV structural proteins were visible, these impurities, possibly cell nuclei and debris, comprised most of the preparation and compromised the visibility of the structural proteins. The presence of these impurities might be due to contaminants with similar densities as the virion particle, which was extracted with the virus band after sucrose gradient purification. Following virus extraction, the sample was concentrated which also resulted in further concentration of the impurities. The number of contaminants present in the sample was clearly more than the virus particle and could therefore possibly explain the difficulty in visualising the viral proteins.

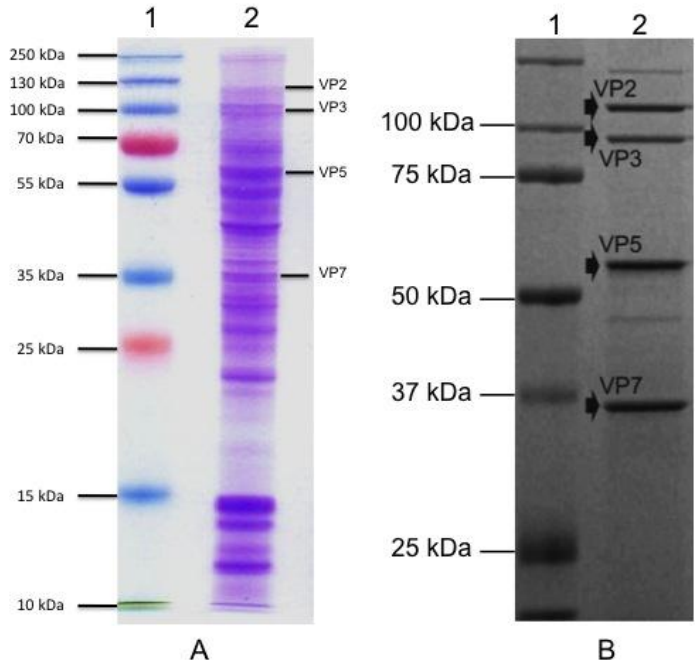

Figure 4.3 SDS-PAGE analysis of the separated structural proteins of AHSV4 after sucrose gradient ultracentrifugation. Figure A lanes: 1) 5 µl PageRuler protein marker; 2) 10 µl of the cultivated AHSV4 indicating the structural proteins VP2, VP3, VP5 and VP7. Figure B illustrates the separated proteins of AHSV4 from the literature. This image is used as a reference point for determining the expected protein positions of AHSV4 in the sample that was partially purified by sucrose gradient ultracentrifugation (Manole *et al*., 2012)

In order to be able to perform proteolytic digestion studies on the proteins of AHSV4, the virus sample had to be purified to an extent in which the viral proteins were more noticeable than the surrounding impurities. By subjecting the partially purified virus to an additional purification step, such as CsCl gradient purification, excess contaminants in the sample could possibly be removed, which would enable better visualisation of the viral proteins.

4.3.3 CsCl ultracentrifugation purification of AHSV4

In order to further purify the AHSV4 sample, a high speed, long duration ultracentrifugation step was employed using a CsCl gradient. Since AHSV4 has a buoyancy of around 1,4 g/cm^3 in CsCl (Roy, 2011), a CsCl-virus mixture was prepared to a refractive index of around 1,4 g/cm^3 and ultracentrifuged, followed by collection of the virus band (Figure 4.4) and visualisation of the viral proteins on a SDS-PAGE gel.

Following centrifugation, two distinct bands were visible under slit light in the CsCl suspension (Figure 4.4). The top band that formed was yellowish in colour and mostly comprised of lipids, cell debri and other impurities. The presence of these impurities was also a clear indication that sucrose gradient ultracentrifugation was rather ineffective. This could explain the abundance of impurities seen on the SDS-PAGE gel in Figure 4.3. The second band visible had the distinct blue-grey colour of viruses when viewed under light (Arnold *et al.*, 2009). The possible virus band was extracted from the gradient solution and the density of the band was determined using a refractometer. The band had a density of around 1,41 g/cm^3 and since this was the approximate density of AHSV4 (Roy, 2011), this band was perceived to be that of AHSV4.

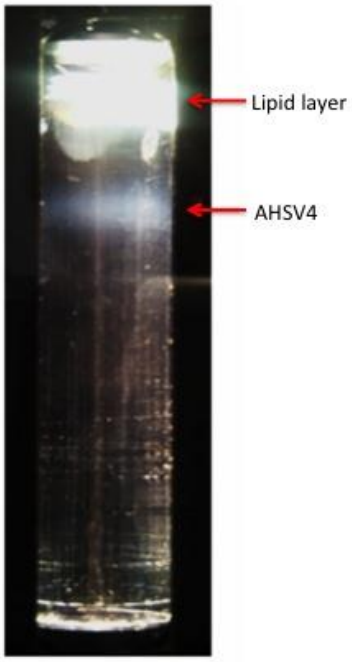

Figure 4.4 Sucrose gradient partially purified AHSV4 band after CsCl gradient ultracentrufigation. Impurities present in the partially purified sucrose gradient AHSV4 preparation were separated from the impurites during ultracentrifugation and the virus band that formed at a density of around 1,4 g/cm^3 was visualised under slit light.

Proteins of the CsCl purified AHSV4 were separated on a SDS-PAGE gel and visualised with Coomassie brilliant blue staining (Figure 4.5 A). Distinct protein bands were visible at the expected molecular masses for AHSV viral proteins,

namely 120 kDa, 103 kDa, 59 kDa and 38 kDa. These bands correlated with expected sizes of the structural proteins of AHSV4 (Figure 4.5 B). Additional protein bands were still visible and could be attributed to impurities as a result of incomplete purification or contamination of the virus preparation during the extraction of the virus band. As the proteins of interest were visible and only a partially purified sample was needed for subsequent protease digestion analyses (Palmer, 2001), the possibility of VP2 being digested by recombinant expressed proteases could subsequently be investigated by incubating the recombinant r*CulsonLTRYP* protease with the CsCl gradient purified AHSV preparation.

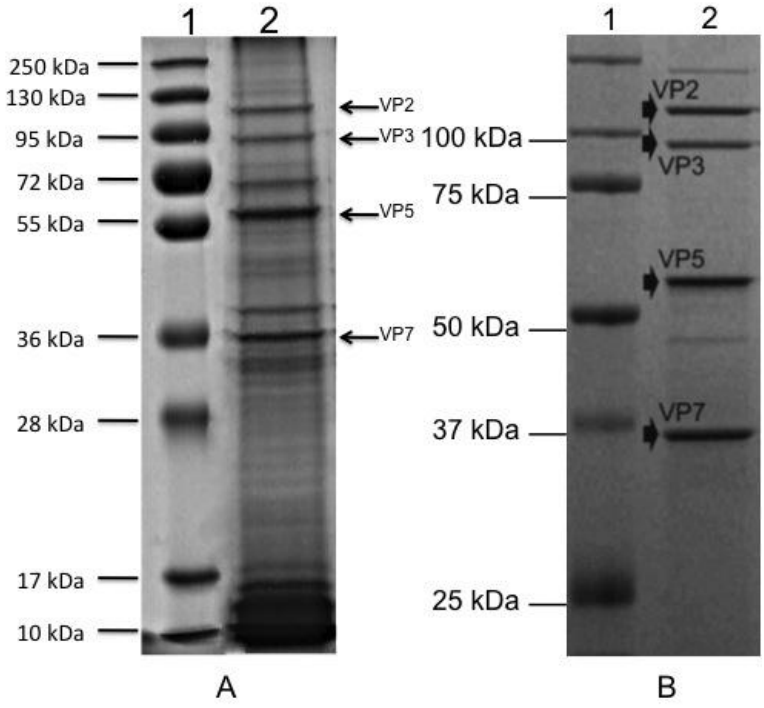

Figure 4.5 SDS-PAGE analysis of the structural proteins of AHSV4 following CsCl ultracentrifugation. Lanes: 1) 5 µl PageRuler protein marker; 2) 10 µl CsCl partially purified AHSV4. Figure B illustrates the separated proteins of AHSV4 that served as a reference point for determining the molecular mass of the separated AHSV structural proteins (Manole *et al.,* 2012)

4.3.4 Digestion of AHSV4 using a *Culicoides* recombinant protease

Darpel *et al* (2011) have reported that proteases present in the saliva of adult *Culicoides* are capable of directly modifying the structure and infectivity of BTV. By incubating the purified salivary proteases with purified BTV for different periods of

time at different temperatures, the structure of BTV VP2 was altered. This resulted in a subsequent increase in the infectivity of BTV in a *C. sonorensis* derived cell line (Darpel *et al*., 2011). It was also stated that these proteases might be capable of altering the structures of other Orbiviruses (Darpel *et al*., 2011). Therefore, the possible effects of recombinant trypsin-like serine protease of *C. sonorensis* on AHSV4 had to be investigated. Since AHSV and BTV are structurally very similar (Howell, 1962), it was decided to try and cleave the structural protein, VP2, of AHSV4 through the use of the expressed recombinant late trypsin-like serine protease, r*CulsonLTRYP* (Chapter 3). Following incubation of the partially purified r*CulsonLTRYP* and CsCl partially purified AHSV4 (1 hour, 3 hours and 6 hours), denatured samples were separated by SDS-PAGE and visualised by means of silver staining (Figure 4.6)

When comparing the incubated VP2 and VP3 with the control sample, the incubated VP2 and VP3 were nearly invisible. This could be attributed to two possible factors. Firstly, it was speculated that the addition of the recombinant protease to the purified virus sample might have resulted in the dilution of the virus sample. Since the concentration of these proteins were low to start with, this dilution might have been to such and extent that the concentration of VP2 and VP3 were so low making it nearly impossible to visualise the proteins. But, when comparing the incubated samples (Lanes 3-5) to the virus control (Lane 2), VP2 and VP3 of the control appear to be more prominent than those of the incubated samples. The intensity of VP5 and VP7 also seemed to have remained unchanged when incubated in the presence of r*CulsonLTRYP*. Based on these findings, the poor visibility of VP2 and VP3 in Figure 4.6 A, lanes 3-5, was not as a result of dilution.

The second possibility for the low visibility of VP2 and VP3 is the complete digestion of these proteins by the recombinant protease. As mentioned earlier, the visibility of VP5 and VP7 (Lanes 3-5) was identical to the control (Lane 2), but VP2 and VP3 of the incubated samples were of a lower intensity than the control sample. The main aim of this part of this experiment was to determine if VP2 could be cleaved by a recombinant trypsin-like *Culicoides* protease. The expressed protease was established to be primarily trypsin-like (Chapter 3), but incomplete

purification has resulted in other proteases also being present in the sample. These impurities, could possibly have contributed to the digestion of VP2 and VP3. Even though it was established that the unknown cellular protein (Figure 4.6, B) co-eluted with r*CulsonLTRYP* is not a protease, its function is not yet established and could therefore also have contributed to the digestion of VP2 and VP3. Hence, the decrease in VP2 and VP3 visibility after incubation with the trypsin-like proteases was probably as a result of protein digestion. The digestion of BTV with the salivary proteins of *C.sonorensis* and trypsin resulted in VP2 of BTV being digested and formed two distinct smaller protein bands (Darpel *et al*., 2011). Even though the AHSV4 samples were incubated for different periods of time with the recombinant protease, no significant digestion of VP2 or VP3 occurred that could be compared to the protein digestion patterns achieved by Darpel *et al* (2011). Therefore, based on the work performed in this chapter, the ability of recombinant *C. sonorensis* proteases to digest AHSV4 VP2 is still undetermined.

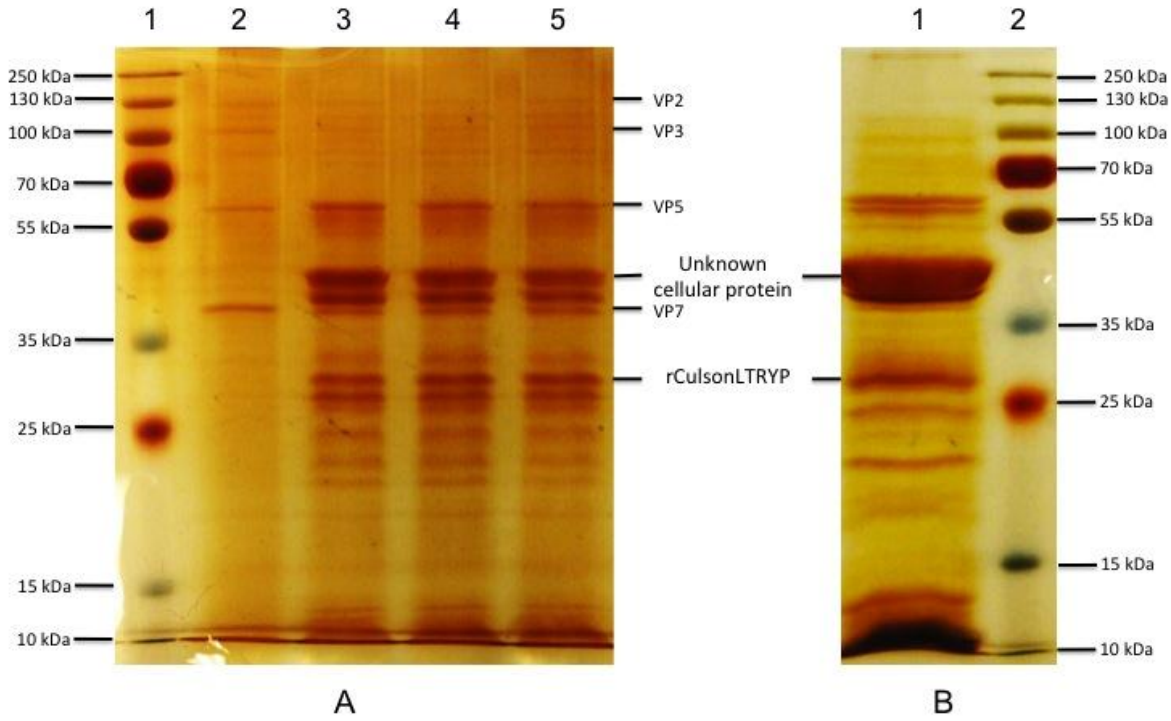

Figure 4.6 Silver stained SDS-PAGE analysis of AHSV4 incubated with r*CulsonLTRYP*. Figure A lanes: 1) 5 µl PageRuler protein marker; 2) AHSV4 control; 3) 1 hour incubation; 4) 3 hour incubation; 5) 6 hour incubation. Incubated samples contained 20 µl of the CsCl partially purified AHSV4 and 10 µl purified r*CulsonLTRYP*. The control sample consisted of 20 µl purified AHSV4 and 10 µl LTB. Figure B lanes: 1) Partially purified and expressed r*CulsonLTRYP* 2) 5 µl PageRuler protein marker. Figure 4.6 B is represented as a control for the expressed recombinant *CulsonLTRYP* incubated with AHSV4 as shown in Figure 4.6 A.

4.4 Summary

Darpel *et al* (2011) have reported that the proteases in the saliva of *C. sonorensis* cleave VP2 of BTV, which result in an increase in the infectivity of the virus. Interestingly, Manole *et al* (2012) determined the structure of a naturally occurring mutant AHSV strain in which VP2 is truncated and possibly structurally similar to the proteases digested VP2 of BTV. This AHSV7 with the truncated VP2 also had a higher infectivity in mammalian cells than the original AHSV7 strain (Manole *et al.,* 2012).

The two main aims of this part of the study were to first cultivate and purify AHSV4 and to determine if a recombinant salivary protease from *C. sonorensis* was

capable of digesting VP2 of AHSV4. The infective virus was used to infect BHK-21 cells, and was subsequently used to cultivate the virus. The virus was partially purified through two ultracentrifugation steps namely sucrose gradient ultracentrifugation and CsCl ultracentrifugation. The partially purified virus was incubated with the purified recombinant protease, r*CulsonLTRYP* (Chapter 3), and the possible effects of virus-protease incubation was analysed on a silver stained SDS-PAGE gel. The process of purifying AHSV4 through the use of only sucrose gradient ultracentrifugation yielded a virus sample that was only partially pure. As seen on the SDS-PAGE analysis in Figure 4.5 A, the purification process was only somewhat successful, since several impurities were still present. Amongst these impurities, structural proteins, possibly VP2, VP3, VP5 and VP7 of AHSV4 could be seen. In order to perform protein digestion studies, the virus sample had to be of much higher purity. Therefore, the sample was subjected to an additional purification step namely CsCl ultracentrifugation. This enabled further separation of the virus from the impurities. The blue-grey virus band that formed was retrieved and the buoyancy was measured to be 1,41 g/cm^3, well within the buoyant density range for AHSV (Roy, 2011).

There was very little AHSV4 recorded after the CsCl gradient ultracentrifugation. The AHSV4 preparation was very dilute, which subsequently resulted in poor visibility of the viral structural proteins on a SDS-PAGE gel. The intensity of the VP5 and VP7 protein bands seemed to be unaffected after incubation with the recombinant protease. However, the intensity of the VP2 and VP3 proteins seemed to be much lower when compared to the undigested control sample. In Chapter 3 it was established that the partially purified r*CulsonLTRYP* preparation had some characteristics of aspartic proteases. Also, r*CulsonLTRYP* was partially inhibited by TPCK, an inhibitor known for its weak ability in inhibiting trypsin (Salvesen & Nagase, 2001). The impurities surrounding r*CulsonLTRYP* might have attributed to these characteristics, which contributed to the visibly lowering of VP2 and VP3's concentration following digestion. The virus was also incubated with the recombinant protease for different periods of time but no visible changes in digestion patterns occurred. As a result of the low virus concentration, it was rather difficult to accurately determine whether or not VP2 and VP3 were digested

by the recombinant late trypsin-like serine protease of *C. sonorensis* or by other proteases in the sample preparation.

Some ways in which these obstacles could be overcome is by concentrating the purified virus and removing the proteolytic impurities in the recombinant protease sample through an optimised purification system. Darpel *et al* (2011) also stated that the efficacy of VP2 cleavage is influenced by the enzyme-virus particle ratio. By determining the ideal ratio of purified virus to recombinant protease for digestion, a more definitive digestion pattern could possibly be achieved. This could aid in future investigations based on the cleavage of AHSV4 structural proteins through the use of recombinant proteases.

CHAPTER 5

Determining the nucleotide sequence of amplified *C. imicola* from cDNA

5.1 Introduction

One of the major contributors in the transmission of AHSV and BTV is the biting midge, *C. imicola* (Boorman & Wilkinson, 1983; Nevill *et al*., 1992). It has also been proven that *C. sonorensis* acts as host for both AHSV and BTV (Mellor *et al.*, 1975; Boorman *et al*., 1975). Darpel *et al* (2011) identified a protein in the saliva of *C. sonorensis* that cleaves the outer capsid protein, VP2, of BTV resulting in an enhanced infectivity of the virus. A protein in *C. imicola* that showed similar proteolytic activity to the 29 kDa protein of *C. sonorensis* was also observed in Chapter 2. Since very little is known about the proteins of *C. imicola*, it was decided to try and identify and sequence this specific protein.

By studying similar proteins in the vectors for AHSV and BTV, comparisons of the viral proteins can be made, which in return could help in the identification of their functions. The characterisation of these proteins could also help in better understanding how these specific proteins assemble into larger complexes and the underlying mechanism involved in increasing infectivity. The composition of the proteins in the saliva of the insect vector for BTV and possibly AHSV, *C. sonorensis*, is directly related to the transmission efficiency (Marchi *et al.,* 1995; Darpel *et al.*, 2011). Therefore, the identification of similar proteins in the insect vector for AHSV, *C. imicola*, would be of importance in better understanding these proteins, and the manner in which they assist in the transmission of Arboviruses.

To summarise, the main aims of this part of the study was to amplify *C. imicola* DNA using *C. sonorensis CsLTRYP3A* primers, after which the amplicon would be sequenced and analysed using web-based software.

5.2 Materials and Methods

5.2.1 *C. sonorensis* late-trypsin *CsLTRYP3A* BLAST

As stated in earlier chapters, the sequence of the late trypsin-like serine protease of *C. sonorensis* (*CsLTRYP3A*) is published (Darpel *et al*., 2011) and available at GenBank (Accession number: AY603563.1). The basic local alignment software tool (BLAST) software available on the website of National Center for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov) was used to collect sequences showing substantial identity to the *CsLTRYP3A* protein from the GenBank database. The *CsLTRYP3A* nucleotide sequence was submitted, and the program selected was optimised for highly similar sequences. Two sequence homologues were identified and the sequences downloaded in FASTA format.

5.2.2 Primer design

Primers were designed (Table 5.1, 5.2 and 5.3) on the outside of the reading frames of the sequence homologues (Sequence 5.1, 5.2 and 5.3) identified as described in section 5.2.1.

>*C. sonorensis* late trypsin- *CsLTRYP3A*

GATCAGTATCATTTTAAAGGAGGAAGTTTGAA**GCTTCACTTTAAAAATGCAG**TTATTTCA
AGTAATTTTCATTTTGATTACATTTTCCTTCATCGAGATATTTGGTGATGAGGTTCAAGA
AATTGTTCCAGTTCCACCTTTAAATTACACTAAACCACCTTACAGTGTAAAAATTGTTGG
TGGTTCTCCTGCCCGCTTGCATCAATTTCCCTGGCAAGCATCTATCACGTCATGCGAAGG
AGGAAGTTGTTATATTTGTGGTGGTTCTTTGATCTCGAAGCGATATGTTTTGACAGCAGC
GCATTGCGCAGCCGGTTTAACCCGATTTGTCATTGGTCTTGGTTCCAATTCGAGAAATCG
ACCTGCTGTTACGTTAACTTCAAATATAAAAGTTGTACACCCACAATACGATGCTAAAAG
CTTAGGAAATGATGTTGCTGTCATTAAATTACCCTGGAGTGTGAAATTAAACAAGGCAAT
TCAACCAATCATTTTGCCACGCTCAAATAATACATACGATAATGCAAATGCAACTGTATC
GGGATATGGGAAAACAAGCGCTTGGAGTTCGTCATCTGATCAATTGAATTTTGTTGATAT
GCGTATTATATCAAATGGTCAATGTAGAGAAATATTTGGATCAGTAATTCGTGATTCGAG
TTTGTGTGCTGTAGGAAAAAATTTGTCAAGGCAAAATGTATGTCAAGGTGATTCTGGAGG
TCCATTGGTGGTTAAAGAAGGAAATTCAACTGTCCAAGTCGGAGTTGTATCTTTTGTATC
TGCAGCTGGATGTGCAGCTGGTTATCCATCTGGATATGCTAGAGTTAGCTCATTTTATGA
ATGGATTGCCAATATGACAGACATT**GATTTGTAATGAAGAAAAATGAAT**AAAAATTTACA
TACCTTGAGTAGCCCAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 5.1 Nucleotide sequence of late trypsin-like serine protease *CsLTRYP3A* of *C. sonorensis*. The position of the three different forward primer sequences are indicated at the 5' end in blue, bold and underligned. The position of the reverse primers are indicated at the 3' end in bold, green and underligned. The start and stop codons are indicated in red.

Table 5.1 *CsLTRYP3A* oligonucleotide primers used for PCR reactions

Primer Name	Primer Sequence	Primer Length (bp)	GC (%)	Melting Temperature (°C)
*Culson*A_F1	GCTTCACTTTAAAAATGCAG	20	35.0	53.1
*Culson*A_F2	TTAAAGGAGGAAGTTTGAAGC	21	38.1	55.5
*Culson*A_F3	CAGTATCATTTTAAAGGAGGAAGTTTGA AGCTTCACTTT	39	33.3	70.3
*Culson*A_R1	ATTCATTTTTCTTCATTACAAATC	24	20.8	53.7
*Culson*A_R2	GTAAATTTTTATTCATTTTTCTTC	24	16.7	51.0
*Culson*A_R3	CAAGGTATGTAAATTTTTATTCATTTTTC	29	20.7	58.2

>*C. sonorensis* Late Trypsin- *CsLTRYP3B*

GATCATTTTATCAAAGTGGTTCTAAGTTGAAATTT**CAATTTCACTTTAAAAATGC**AGTTA
TTTAAAGTAACTTTCATTTTGATTGCACTTTCCATCATCGAGATATTTGGTGATGAGCTT
CAAGAAATTGTTGCAGTTCCACCTTACAATTACACTAAACCACCTTACAGTGTTAAAATT
GTTGGTGGTTCTCCTGCTCGCGTGCATCAATTTCCCTGGCAAGCATCTATCACGTCATGC
GATGGAGGAAGTTGTTATATTTGTGGTGGTTCTTTGATTTCGAAGCGATATGTTTTGACA
GCAGCGCATTGCGCAGCCGGTTTGACACGATTTATCATTGGTCTTGGATCAAATTCGAGA
AATCGACCTGCAATTACGTTAACTTCAAATATAAAAGTTGTACACCCACAATACGATGCT
AAAAGCTTAGGAAATGATGTTGCTGTCATTAAATTACCCTGGAGTGTGAAATCAAACAAG
GCAATTCAACCAATCATTTTGCCACGCTCAAATAATACTTACGATAATGCAAATGCAACT
GTATCGGGATATGGGAAAACAAGCGCTTGGAGTTCGTCATCTGATCAATTGAATTTTGTT
GATATGCGTATTATTTCAAATAGTAAATGTAGAGAAATATTTGGATCGGTAATTCGTGAT
TCGAGTTTATGTGCTGTAGGAAAAAATCGGTCAAGGCAAAATGTATGTCGAGGTGATTCT
GGAGGTCCATTGGTGGTTAAAGAAGGAAATTCAACTGTCCAAGTCGGAGTTGTATCTTTT
GTATCTGCAGCTGGATGTGCAGCTGGTTATCCATCTGGATATGCTAGAGTTAGCTCATTT
TATGAATGGATTGCCAATATGACAGACATTGATTTG**TAA**TGAAAAAAAATG**AATAAAAAT**
TTACATACCTTGAGTAGCTCAMAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 5.2 Nucleotide sequence of late trypsin-like serine protease *CsLTRYP3B* of *C. sonorensis*. The position of the three different forward primer sequences are indicated at the 5' end in blue, bold and underligned. The position of the reverse primers are indicated at the 3' end in bold, green and underligned. The start and stop codons are indicated in red.

Table 5.2 *CsLTRYP3B* oligonucleotide primers used for PCR reactions

Primer Name	Primer Sequence	Primer Length (bp)	GC (%)	Melting Temperature (°C)
*CulsonB*_F1	CAATTTCACTTTAAAAATGC	20	25.0	50.5
*CulsonB*_F2	CAAAGTGGTTCTAAGTTGA	19	36.8	48.4
*CulsonB*_F3	GATCATTTTATCAAAGTGGTTCTAAGTTG	21	31.0	60.8
*CulsonB*_R1	GCTACTCAAGGTATGTAAATTTTTATTC	28	28.6	56.7
*CulsonB*_R2	AATTTTTATTCATTTTTTTTCATTAC	26	11.5	53.9
*CulsonB*_R3	CATTTTTTTTCATTACAAATCAATGTCTGTC	31	25.8	63.8

> *Culicoides nubeculosus* clone *CNSG15* trypsin mRNA

GATCATTTCGAATTTTACGTTTAAATAATAAA**ATAAAATTCTCTTTAAAATGAAG**TCTTT
TAAATTAATCTTCATTTCTATTGTGTTCTCAGTCTTGGAAGTTCATGGAGGTGAAGTTCA
AGAGATTATGCCAGTTCCTCCTCAAAATTACACCAGAAAACGTATCACTATGAAAATTGT
TGGTGGTTCTCCTGCTAGAGAACATCAATTTCCATGGCAAGCTTCTGTCACTTCATGTTC
TGGACAAAGTTGTTCCACTTGTGGTGGTTCCCTTATTTCAAGAAGATACATTTTGACTGC
AGCTCATTGCACAGCTGGTTTATCTGAATTCCAAATTGGACTTGGAACAAATAAGAGAAA
TAAACCAACTGTAAATTTAATCTCAAAAGTAAAAATTGAACATCCAAAATATCATGCTCG
ATGGTTAATTAATGATGTATCTATTATTAAATTACCAAGAGCTGTACGATTAAATAAAGG
TGTTCAAATAATCCCGTTGGCACGTAAAAATGAGACATTTAATAATCGATATGCAACGGC
ATCAGGATATGGAAAAACAAGCTCTTTGAGTTCGGTATCTGATCAATTAAATTATGTTGA
TATGCGTATTATTTCTAATACGGAATGTAAAAAAACCTATGGATTGTCTGTTCAAGATTC
TACTTTGTGTGCTGTAGGAAAAAATCGGGCAAGCCAAAATGTATGTCAAGGTGATTCTGG
AGGTCCATTGGTAGTTAAAGAGGGAAATTCATTTGTTCAAGTCGGTGTTGTATCTTTTGT
AGCAAAGGCCGGATGTGAGTATGGATTTCCATCTGGATATGCGAGAGTAAGCTCATTTTA
TCCATGGATTATAAGAACTGCTAGTAGACTTTTG**TAA**TTATCATATAAAATAT**GAAAAAT**
AAACGGATTTGGGCCAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 5.3 Nucleotide sequence of the *C. nubeculosus* trypsin clone, *CNSG15*. The position of the two different forward primer sequences are indicated at the 5' end in bold and underligned. The position of the reverse primers are indicated at the 3' end in bold and underligned. The start and stop codons are indicated in red.

Table 5.3 *CNSG15* oligonucleotide primers used for PCR reactions

Primer Name	Primer Sequence	Primer Length (bp)	GC (%)	Melting Temperature (°C)
*Culnub*_F1	ATAAAATTCTCTTTAAAATGAAG	23bp	17.4	49.2°C
*Culnub*_F2	CGTTTAAATAATAAAATAAAATTC	24bp	12.5	49.4°C
*Culnub*_R1	GGCCCAAATCCGTTTATTTTTC	22bp	40.9	62.7°C
*Culnub*_R2	GATAATTACAAAAGTCTACTAGCA	24bp	29.2	49.9°C

The primers were purchased from Integrated DNA Technologies (IDT), Iowa, United States of America. A stock solution of the primers was prepared by dissolving the lyophilised DNA in dH_2O to a final concentration of 100 µM. Working solutions were prepared by diluting the stock solution to a final concentration of 10 µM.

5.2.3 Isolation of *C. imicola* total RNA from midge homogenate

Homogenisation of *C. imicola*

To be able to analyse *C. imicola* at the molecular level, all fractions of the sample had to be equal in composition and the overall molecular make-up of the sample homogenous. This was done with the use of the Qiagen tissue lyser II (Qiagen, Limburg, Netherlands) at ARC-OVI.

Midges were collected and sorted as described in Section 2.2.1. Following sorting, around 500 midges were suspended in 500 µl Tri-reagent to maintain the integrity of the RNA and inhibit any form of RNase activity during homogenisation. The midge solution was transferred to a 1,5 ml microfuge tube and two stainless steel beads (5 mm) were added. The midges were completely disrupted through high-speed shaking at 25 Hz for 10 minutes after which the sample was frozen at -20°C.

C. imicola total RNA isolation

The total RNA of *C. imicola* was isolated using an adapted version the QIAzol protocol for RNA isolation (http://agtc.wayne.edu/pdfs/qiazol_handbook_1_.pdf). Upon receipt of the midge homogenate, the sample was thawed on ice. Once thawed, the homogenate was subjected to additional homogenisation with 30 strokes through an 18 G needle with a 1 ml syringe. Following homogenisation, cellular debri and large particles were collected by centrifugation (16 000 *g*, 10 minutes, 4°C). The supernatant was carefully aspirated to avoid disruption of the pellet and divided between two 1,5 ml tubes. The proteins were denatured to become soluble in the organic phase through the addition of half the volume chloroform to the transferred supernatant. The chloroform-homogenate mixture was vortexed for approximately 30 seconds and incubated at room temperature for about 5 minutes. To avoid RNA degradation, all steps preceding incubation were performed at 4°C. Following incubation, the organic and aqueous phases were separated by means of centrifugation (16 000 *g*, 30 minutes). The upper aqueous phase containing the nucleic acids was transferred to a new tube and half the

volume isopropanol was added to the total volume of Tri-reagent used. The tube was inverted 10 times to ensure thorough mixing after which the homogenate was incubated at room temperature for about 5 minutes. Following incubation, the RNA in the homogenate was pelleted (16 000 *g*, 30 minutes). The supernatant was carefully aspirated without disrupting the RNA pellet. The RNA pellet was resuspended in 1 ml ice-cold 75% (v/v) ethanol and washed by means of vortexing for 30 seconds. Following washing, the RNA was collected by centrifugation (16 000 *g*, 30 minutes). The supernatant was removed completely and the RNA dried by inverting the tube and incubating at room temperature for 30 minutes. The RNA pellet was dissolved in 50 µl RNase-free water. The total RNA from ± 500 *C. imicola* midges was spectrophotometrically quantified on the Nanodrop instrument as described in Section 3.2.6 with the paramenters adjusted to RNA quantification.

5.2.4 cDNA synthesis from *C. imicola* isolated RNA

cDNA from the single stranded RNA isolated from *C. imicola* was synthesised using the Avian Myeloblastosis Virus (AMV)-Reverse Transcriptase (RT) system. AMV-RT is a RNA-directed DNA polymerase capable of synthesizing a complementary DNA strand initiated from a primer using the RNA as a template (Sambrook & Russel, 2001).

The reaction mixture for the synthesis of *C. imicola* cDNA was made up as follows: in a sterile, nuclease-free microcentrifuge tube, Oligo(dT)18 primer (0,5 µg/µl) was added to single stranded RNA (1 µg/µl) and the final volume adjusted to 13 µl with nuclease-free H_2O. The mixture was vortexed briefly to prepare a homogenous mixture. The primer was annealed to the template by heating the mixture at 65°C for 5 minutes followed by chilling on icy water for 10 minutes. In specific order, 5x AMV-RT buffer (4 µl), RNasin® ribonuclease inhibitor (0,5 µl), dNTP mix (2 µl) and AMV reverse transcriptase (0,5 µl) was added to the annealed primer-template mixture to make up a final reaction volume of 20 µl. First strand cDNA was synthesised by incubating the reaction mixture at 50°C overnight after which the enzyme was inactivated at 85°C for 5 minutes. The synthesised cDNA was stored at 4°C until further use.

5.2.5 PCR amplification of *C. imicola* cDNA

The enzyme, DNA polymerase, has the ability to synthesize a new strand of DNA complementary to the template strand through the process of PCR (Saiki *et al*., 1988). In short, primers are used by DNA polymerase to produce the first nucleotide onto the preexisting 3'-OH group. This makes it possible to define the specific region of the template sequence that needs to be amplified. This process enables the generation of billions of copies of the specific sequence of interest.

The widely used enzyme, Ex*Taq* DNA polymerase was used to perform PCR, since it preferentially adds adenine to the 3' end of the products (Sambrook & Russel, 2001). The PCR reaction mixture consisted of approximately 250 ng cDNA, 25 pmol of each primer (Section 5.2.1), 0,4 µl of each dNTP, 1x Takara Ex*Taq* polymerase buffer and 2 units Takara Ex*Taq* DNA polymerase. The final volume of the reaction mixture was adjusted to 25 µl using dH_2O. Unless otherwise stated, the thermal cycling conditions were 94°C for 5 minutes, then 35 cycles of 94°C for 30 seconds, 50°C for 90 seconds, and 72°C for 1 minute. This was followed by a final extension phase of 30 minutes at 72°C. All reactions were performed on a Bio-Rad thermal cycler.

The amplified products were separated on a 1% (m/v) agarose gel and visualised as described in Section 3.2.8. Following this, the amplicons of interest were excised from the gel and purified according to the method described in Section 3.2.9.

5.2.6 TA cloning of amplified *C. imicola* cDNA

Since the primers for amplifying a possible protease in *C. imicola* were long and did not contain any restriction sites, it was decided to use a TA cloning system. By using this technique, the amplified PCR products could be cloned into the TA cloning vector, which could furthermore contribute to the production of significant quantities of amplified DNA (Marchuk *et al.,* 1991). *Taq* polymerase creates dA overhangs at the 3' end of the amplified products, which corresponds to the 3' dT

overhangs of the TA cloning vector. By introducing DNA ligase to the insert and the vector, a recombinant plasmid containing the amplicon sequence will result.

TA cloning was performed using the pTZ57R/T vector. The vector and purified PCR product were combined in a 1:2 molar ratio. A 25 µl ligation reaction contained 1x T4 DNA ligation buffer, 5 Weiss units of T4 DNA ligase, 30 pmol of PCR amplicon and 15 pmol of vector. The ligation was performed in a Bio-Rad thermal cycler at 22°C for 3 hours after which the temperature was reduced to 4°C for 24 hours.

A volume of 15 µl of the ligation reaction mixture was used to transform chemical competent *E. coli* JM109 cells. Transformation and plating were performed as described in Section 3.2.11. Positive colonies were screened using the primers that successfully amplified the DNA at the size of interest. The amplified DNA was visualised on a 1% (w/v) agarose gel.

5.2.7 Amplicon sequence determination and analysis

The recombinant plasmids of the colonies that contained the insert of interest were purified as described in Section 3.2.5. The nucleic acid sequence was determined, at the DNA sequencing laboratory, Central analytical facility, University of Stellenbosch**.** The sequences were determined using Sanger sequencing. FinchTV version 1.40 (www.geospiza.com/finchtv) was used to analyse the chromatograms and the sequence was aligned with reference sequences using DNAMAN version 6 software. The open reading frames of the sequences were also identified using the DNAMAN software. The sequences obtained were subjected to the NCBI BLAST tool to determine if there were any sequence homologies between the sequences obtained and the GenBank database. The open reading frame of the nucleotide sequence was translated using the ExPasy translator tool (www.web.expasy.org/translate). In order to determine if the amplified product was a protease, the protein sequence was submitted to an online protease identifier tool (www.csbio.sjtu.edu.cn/bioinf/protease).

5.3 Results and Discussion

5.3.1 *C. sonorensis* late-trypsin *CsLTRYP3A* BLAST

The ability of *C. sonorensis* and *C. imicola* to act as a host for both AHSV and BTV (Boorman *et al*., 1975; Mellor *et al.*, 1975) has given the opportunity to investigate possible similarities in the ability of these vectors to increase infectivity of AHSV and BTV. It is known that a trypsin-like serine protease present in the saliva of the BTV vector, *C. sonorensis*, cleaves VP2 of the virus, resulting in increased virulence (Darpel *et al.*, 2011; Marchi *et al*., 1995). Little is known about the manner in which *C. imicola* transmits AHSV, but it can be postulated that *C. imicola* and *C. sonorensis* use similar proteins in the transmission process. In order to determine if there are any other known proteins that are similar in function, the sequence of the late trypsin-like serine protease of *C. sonorensis* was used as the query sequence against all known sequences in the NCBI database. The BLAST tool was used to determine if any sequence identities were present.

As seen in Figure 5.4, the BLAST tool identified two sequences similar to the *C. sonorensis* late trypsin-like serine protease, *CsLTRYP3A*. The sequence with the highest identity to *CsLTRYP3A* was also from *C. sonorensis*, namely late trypsin *CsLTRYP3B* (AY63564.1). This sequence was also trypsin-like and had a 96% identity to the query sequence. The second sequence identified to be similar to the query sequence was that of *C. nubeculosus*. The *CnSG15* (EU978886.1) sequence had a 76% identity to *CsLTRYP3A*. *C. nubeculosus* is the vector for parasitic infection, *Onchocerca cervicalis*, in horses (Smidt *et al.*, 1982). The bite of the vector hosting this parasite causes severe dermal allergic reactions in horses. Furthermore, previous laboratory studies performed on this vector have also shown that *C. nubeculosus* is capable of being a host to BTV (Boorman, 1974).

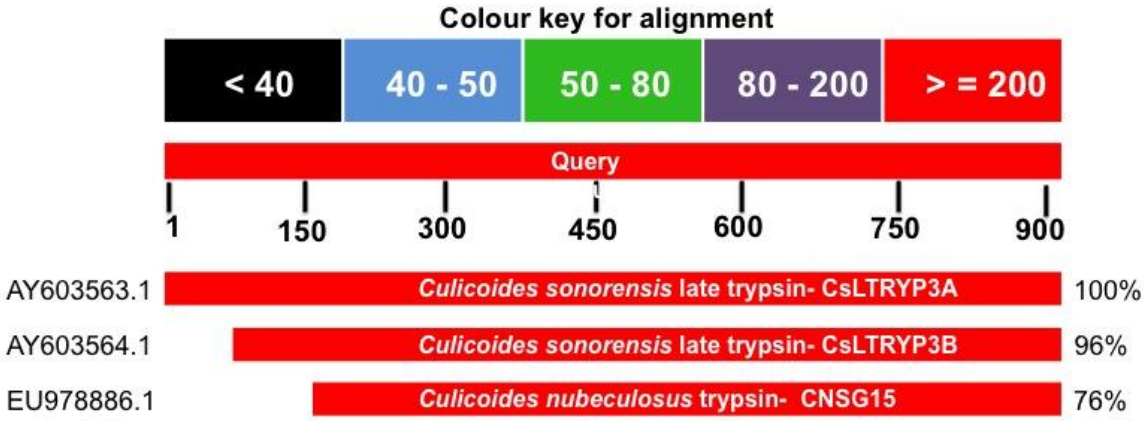

Figure 5.4 Sequence homologies obtained by the NCBI BLAST tool. The colour key is used to indicate the *Culicoides* protease intensity of the sequence identity, with black indicating few similarities and sequences very similar to the query sequence in red. The accession number of the identified sequences is indicated on the left-hand side of the identified sequence and the similarity percentage on the right. Two sequences were identified to be similar to the *C. sonorensis* late trypsin *CsLTRYP3A* query sequence. The *C. sonorensis* late trypsin-like protease, *CsLTRYP3B* has shown an identity match of 96% to the query sequence. The vector for BTV, *C. nubeculosus*, also had a protein which had a 76% identity to that of the query sequence.

Two different vectors for BTV have been shown to have a protease very similar to the sequence of the *C. sonorensis* late trypsin-like serine protease A (Darpel *et al*., 2011; Boorman, 1993). These proteases are responsible for cleaving VP2 of BTV that ultimately results in an increased infectivity of the virus (Van Dijk & Huismanns, 1982; Darpel *et al.,* 2011). Based on these findings, it was postulated that the vector that act as host for AHSV and BTV, *C. imicola*, had a protein of similar function. To verify this theory, a nucleotide sequence similar in size to the late trypsin-like serine protease, *CsLTRYP3A*, had to be identified in *C. imicola*.

5.3.2 Amplification of *C. imicola* cDNA

In order to determine if *C. imicola* had a protein similar to the proteins identified in *C. sonorensis* and *C. nubeculosus*, primers had to be designed based on the sequences identified in Section 5.3.1. Different sets of primers were designed on the outside of the open reading frames. A total of three sets forward and reverse primers were designed for the nucleotide sequences of *CsLTRYP3A* and *CsLTRYP3B* and two forward and reverse primers for *CNSG15*. Upon receipt, the primers were rehydrated with dH_2O to 100 µM and diluted to a working

concentration of 10 µM. *C. imicola* midges were collected and sorted, the RNA isolated and cDNA was synthesized using the isolated RNA. The *C. sonorensis* trypsin-like serine protease has an open reading frame of around 830 nucleotides. Therefore, an amplicon with a size of around 830 bp could possibly suggest that *C. imicola* has a protein similar to the *CsLTRYP3A* / *CsLTRYP3B* protein of *C. sonorensis.* The cDNA of *C. imicola* was subjected to PCR using the primer sets designed from the BLAST identified sequences. The outcome of all of the PCR experiments were analysed on a 1% (w/v) agarose gel.

As seen in Figure 5.5, the primers of *Culson*B (Lanes 2-4), *Culson*A_1 (Lane 5) and *Culnub* (Lanes 8-9) were unable to amplify any part of *C. imicola*. This could possibly be explained by the inability of the primers to bind to the template. Since no PCR program with optimised parameters for the amplification of *C. imicola* with non-specific primers was available, a general PCR program was used to start with. The annealing and extension times used in this initial PCR experiment were relatively short and not specific to the individual primers used. Therefore, the primers probably didn't have enough time to bind to the template resulting in no visible amplification. An annealing temperature might also have been too high for these specific primers, which could also explain the inability of the primers to bind to the template (Roux, 1995).

Some form of amplification was visible with the use of *Culson*A_3 and *Culson*B_3 primers (Lanes 4 and 7). This amplification could be seen below 100 bp and was probably PCR by-products known as primer dimers. Primer dimers usually occur when the primer molecules hybridise to each other as a result of complimentary bases in the primers. This leads to the amplification of the primer dimer by DNA polymerase and ultimately creates competition for the PCR reagents (Roux, 1995). The ± 100 bp primer dimers (Lanes 4 and 7) were of moderate intensity. The long length of the primers and incorrect primer concentration used might also have contributed to an increased ability of the primers to bind to each other.

Even though a single amplicon of around 830 bp was expected, *Culson*A_2 was able to produce four amplified products (Lane 6). These amplicons were respectively around 2500 bp, 1850 bp, 1300 bp and 950 bp in lengths. An excess

amount of cycles in the PCR program possibly increased the opportunity for non-specific amplification and errors. The initial high concentration of primers used (35 pmol) could also explain the increased ability of the primers to bind to non-specific sites on the template. Furthermore, by optimising the thermal cycling conditions, more specific amplification could possibly be achieved (Roux, 1995).

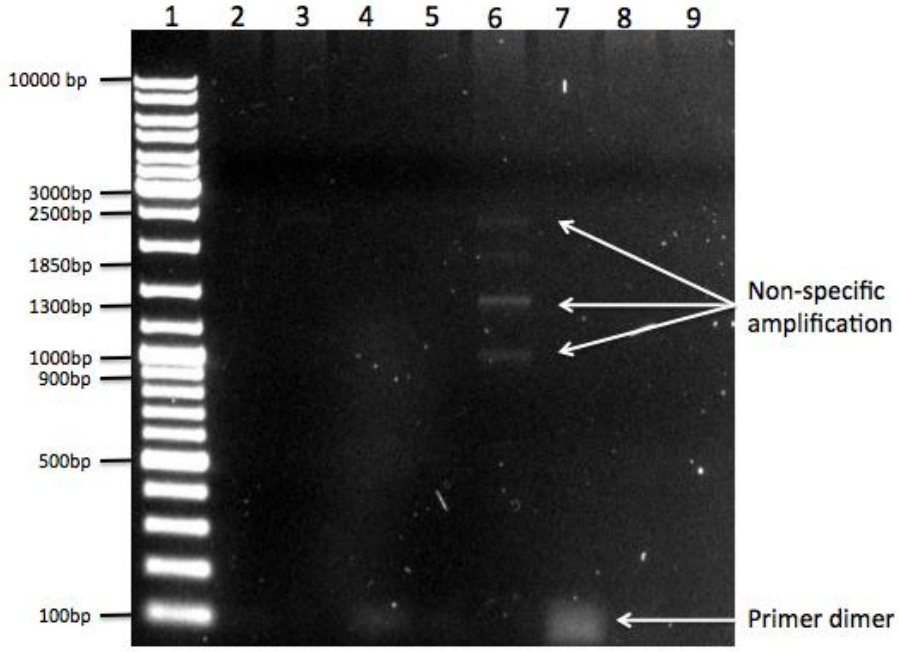

Figure 5.5 Agarose electrophoretic analysis of the PCR amplification of *C. imicola*. Lanes: 1) 5 µl O'GeneRuler DNA marker; 2) *Culson*B_F1 and *Culson*B_R1; 3) *Culson*B_F2 and *Culson*B_R2; 4) *Culson*B_F3 and *Culson*B_R3; 5) *Culson*A_F1 and *Culson*A_R1; 6) *Culson*A_F2 and *Culson*A_R2; 7) *Culson*A_F3 and *Culson*A_R3; 8) *Culnub*_F1 and *Culnub*_R1; 9) *Culnub*_F2 and *Culnub*_R2.

The general rule of thumb in PCR is to use an annealing temperature that is around 5°C lower than the Tm of the primers (Sambrook & Russel, 2001; Roux, 1995). The combined average Tm of the *Culson*A_F2 and *Culson*A_R2 (Lane 6) was 53°C and based on this, the annealing temperature was adjusted 50°C. The primers used was longer than recommended, therefore, the annealing time was increased to 90 seconds. Since long dA overhangs at the 3' end of the amplicon was needed for an upcoming part of the study, the final extension phase was increased from 5 minutes to 30 minutes.

As seen in Figure 5.6, Lane 2, the optimisation of the PCR conditions, yielded more specific amplification. Three bands were amplified at around 1500 bp, 1100 bp and 830 bp. Since the primers were not specifically designed for *C. imicola*, non-specific amplification was expected. The amplicon visible at around 830 bp was similar in size to *CsLTRYP3A* and *CsLTRYP3B*. Therefore, all further analyses were focused on this specific amplicon only.

In order to produce a large population of the amplified DNA, the 830 bp segment and a cloning vector needed to be assembled to produce a recombinant DNA. By producing a large population of the 830 bp amplicon, the amplified DNA could be sequenced, which ultimately could aid in the prediction of the amino acid sequence of the protein that this specific gene encodes. The 830bp amplicon was excised from the agarose gel and purified. The prolonged final extension phase produced long dA overhangs at the 3' ends of the amplicon (Roux, 1995). The pTZ57R/T cloning vector had long 3' dT overhangs that corresponded to the amplicons' dA overhangs which enabled the ligation of the insert to the vector without the use of restriction sites. To confirm that the insert was successfully cloned into the vector, well-isolated transformed colonies were screened by means of colony PCR.

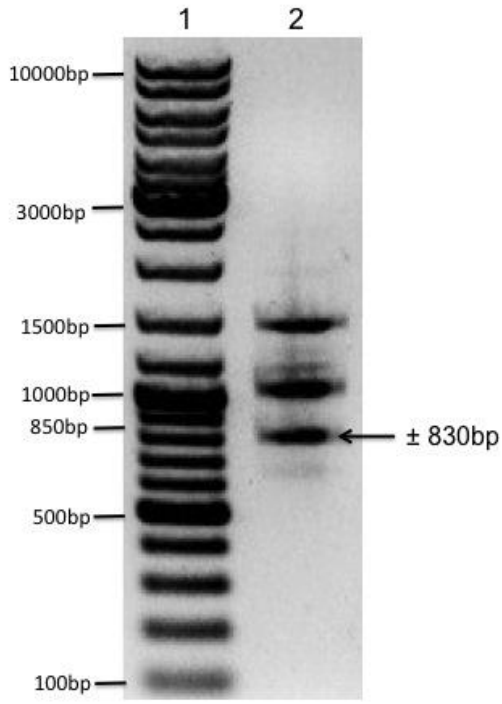

Figure 5.6 Agarose electrophoretic analysis of the optimised PCR amplification of *C. imicola* using *Culson*A_F2 and *Culson*A_R2 primers. Lanes: 1) 5 µl O'GeneRuler DNA marker; 2) *C. imicola* amplified with *Culson*A_F2 and *Culson*A_R2.

As seen in Figure 5.7, the amplified segment of *C. imicola* was successfully cloned into the pTZ57R/T cloning vector. The amplified insert was clearly visible at around 830 bp following colony PCR. This size corresponded to the expected size of the insert, confirming successful ligation. By cloning the amplified 830 bp *C. imicola* band into the TA cloning vector, a large population of the amplified DNA could be generated. This enabled further analyses of the amplicon such as sequence determination.

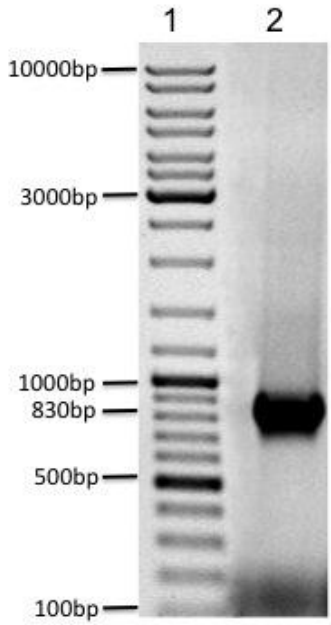

Figure 5.7 Agarose electrophoretic analysis of PCR screening of the cloned 830 bp *C. imicola* amplicon. Lanes: 1) 5 µl O'GeneRuler DNA marker; 2) Amplified 830 bp *C. imicola* cloned into the pTZ57R/T cloning vector.

5.3.3 Sequence analysis of the 830bp *C. imicola* amplicon

In order to determine if the amplified 830 bp fragment in *C. imicola* had sequence identities to the 29 kDa trypsin-like serine protease of *C. sonorensis*, the order of the nucleotides of the amplicon needed to be determined. By determining the *C. imicola* amplicon nucleotide sequence, sequence comparisons could be made between *C. imicola* and *C. sonorensis*. This could help in better understanding the differences and similarities between similar sized genes in these specific species.

As seen in Figure 5.8, the intensities of the signals on the chromatogram were high. The height of these peaks corresponded to the relative concentration of a specific base at that position in the sequence. The high intensity of the peaks was a good indication that high concentration bases were present in the sample. Very few to no baseline peaks were present. This was a good indication that the template and primer was of high quality. Apart from the peaks on the chromatogram, the quality value for the DNA sequence was used to assess the accuracy of each of the bases in the DNA sequence. The software calculated these values by taking the $\log_{10}$ of the error probability and multiplying it by -10.

The higher the quality value obtained, the lower the error probability was. The majority of the sequencing quality values were above 45 indicating an error probability of less than 1 in 10 000.

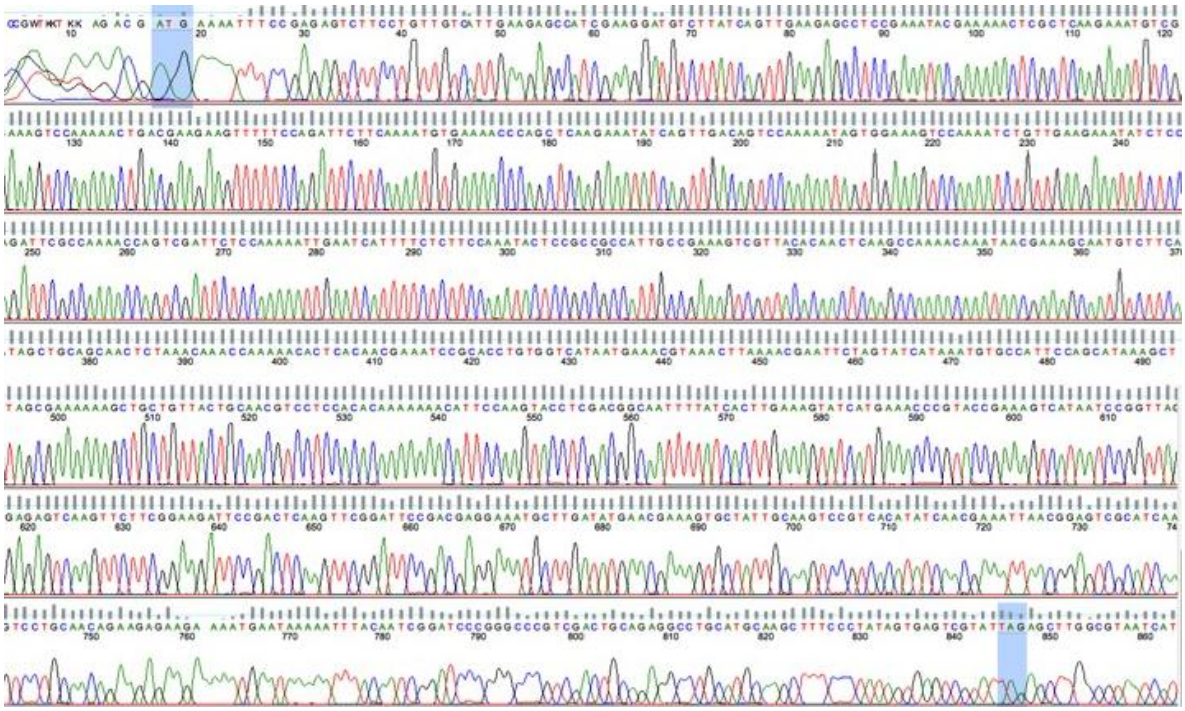

Figure 5.8 Chromatogram of the sequenced *C. imicola* amplicon. The area highlighted in blue at the 5' end indicates the start codon of the sequence and the 3' highlighted area indicated the stop codon. The signal intensities are presented in a graph with the four bases, each identified by different colours. Adenine is indicated in green, thymine in red, cytosine in blue and guanine in black.

Based on the results seen in Figure 5.8, the complete nucleotide sequence was entered into DNAMAN software and the open reading frame of the sequence was determined (Figure 5.9). The amplified *C. imicola* DNA had an open reading frame of 830 bp, which after being translated (www.web.expasy.org/translate) corresponded to 276 residues. When taking into consideration that the molecular weight of water is 18.015 g/mol, the molecular weight of the amplicon could be calculated using a simple equation (Equation 6.1) (www.pir.georgetown.edu).

Equation 5.1 Molecular weight calculation

Molecular weight =
(sum of individual residues weights - water molecular weight) x (number of residues - 1)

The open reading frame of the sequenced amplicon of *C. imicola* had a molecular weight of 29,7 kDa. The most abundant residue was serine with 46 out of the 276 residues. This resulted in 16.7% of the sequence being serine. Proline residues comprised 10.9% of the sequence and lysine 10.5%. Based on the size and molecular weight of the amplified *C. imicola* amplicon, there were some similarities between the late trypsin-like serine protease, *CsLTRYP3A*, (Darpel *et al*., 2011) and the amplified *C. imicola*. Both had a molecular weight of around 29 kDa and comprised of about 830 nucleotides. Since the molecular weight is calculated based on the individual weights of each residue, it was speculated that the residue contents of the two sequences were very similar. However, to confirm this speculation of the *C. imicola* amplicon being similar to the *C. sonorensis* trypsin-like protease, the open reading frame of the *C. imicola* sequence had to be aligned with the open reading frame of *CsLTRYP3A*. This was done using DNAMAN version 6 software. Alignment has shown that the two sequences were in fact only 34.19% similar. Since no trypsin-like protease signature sequences were identified in the sequence alignment between the *C. imicola* amplicon and the sequence of *CsLTRYP3A*, the possibility of the amplified *C. imicola* being a trypsin-like serine protease was eliminated.

```
  1 atgaaaatttccgagagtcttcctgttgtcattgaagagccatcg
    M  K  I  S  E  S  L  P  V  V  I  E  E  P  S
 45 aaggatgtcttatcagttgaagagcctccgaaatacgaaaaactc
    K  D  V  L  S  V  E  E  P  P  K  Y  E  K  L
 90 gctcaagaaatgtcggaaagtccaaaaactgacgaagaagttttt
    A  Q  E  M  S  E  S  P  K  T  D  E  E  V  F
135 ccagattcttcaaaatgtgaaaacccagctcaagaaatatcagtt
    P  D  S  S  K  C  E  N  P  A  Q  E  I  S  V
180 gacagtccaaaaatagtggaaagtccaaaatctgttgaagaaata
    D  S  P  K  I  V  E  S  P  K  S  V  E  E  I
225 tctccagattcgccaaaaccagtcgattctccaaaaattgaatca
    S  P  D  S  P  K  P  V  D  S  P  K  I  E  S
270 ttttctcttccaaatactccgccgccattgccgaaagtcgttaca
    F  S  L  P  N  T  P  P  P  L  P  K  V  V  T
315 caactcaagccaaaacaaataacgaaagcaatgtcttcaatagct
    Q  L  K  P  K  Q  I  T  K  A  M  S  S  I  A
360 gcagcaactctaaacaaaccaaaaacactcacaacgaaatccgca
    A  A  T  L  N  K  P  K  T  L  T  T  K  S  A
405 cctgtggtcataatgaaacgtaaacttaaaacgaattctagtatc
    P  V  V  I  M  K  R  K  L  K  T  N  S  S  I
450 ataaatgtgccattccagcataaagctgtagcgaaaaaagctgct
    I  N  V  P  F  Q  H  K  A  V  A  K  K  A  A
495 gttactgcaacgtcctccacacaaaaaaacattccaagtacctcg
    V  T  A  T  S  S  T  Q  K  N  I  P  S  T  S
540 acggcaattttatcacttgaaagtatcatgaaacccgtaccgaaa
    T  A  I  L  S  L  E  S  I  M  K  P  V  P  K
585 gtcataatccggttaggagagtcaagttcttcggaagattccgac
    V  I  I  R  L  G  E  S  S  S  S  E  D  S  D
630 tcaagttcggattccgacgaggaaatgcttgatatgaacgaaagt
    S  S  S  D  S  D  E  E  M  L  D  M  N  E  S
675 gctattgcaagtccgtcacatatcaacgaaattaacggagtcgca
    A  I  A  S  P  S  H  I  N  E  I  N  G  V  A
720 tcaagtcctgcaacagaagagaagaaaatgaataaaaatttacaa
    S  S  P  A  T  E  E  K  K  M  N  K  N  L  Q
765 tcggatcccgggcccgtcgactgcagaggcctgcatgcaagcttt
    S  D  P  G  P  V  D  C  R  G  L  H  A  S  F
810 ccctatagtgagtcgtattag 830
    P  Y  S  E  S  Y  STOP
```

Figure 5.9 Nucleic acid and protein sequence open reading frame of amplified *C. imicola*. The open reading frame consists of 830 nucleotides and 276 residues. The start and stop codons are indicated in red.

In order to determine if the sequence of the amplified *C. imicola* had any possible sequence identities, the nucleotide (blastn) and protein sequence (blastx) was subjected to BLAST (www.ncbi.nlm.nih.gov) using the default settings of the software. In both of these searches, there were no sequence identities. The algorithm parameters of the search engine were adjusted by increasing the

expected threshold and the filters for low complexity were removed. However, there were still no significant similarities found, meaning that that there are currently no sequences in the database that can relate to the identified *C. imicola* sequence. The possibility of the amplified *C. imicola* to be a protease also had to be investigated. This was done with the use of a web-based protease identifier tool (www.csbio.sjtu.edu.cn/bioinf/protease) that searched for protease domains within the query sequence. However, no proteolytic domains could be identified and it was established that the amplified *C. imicola* was also not a protease.

5.4 Summary

In Chapter 2 a 29 kDa protein in *C. imicola* with trypsin-like characteristics was identified. Furthermore, this protein was similar in size to the trypsin-like serine protease, *CsLTRYP3A*, identified by Darpel *et al* (2011). In Chapter 2 of this study, a protein in the total homogenate of *C. imicola* similar in size to the trypsin-like serine protease of *C. sonorensis* (Darpel *et al*., 2011) was identified. It was also established that this protein had trypsin-like proteolytic activity. The *C. sonorensis* late trypsin-like serine protease is responsible for increasing the infectivity of BTV (Darpel *et al*., 2011). Based on this, *C. imicola* may also use this similarly identified protease for increasing the infectivity of AHSV. Currently the genome of *C. imicola* is not sequenced proving it difficult to identify similarities between the proteins of these vectors. The *C. sonorensis* sequence, *CsLTRYP3A*, is sequenced and published (Darpel *et al.*, 2011). Since these vectors have shown so many similarities, it was decided to use the published *C. sonorensis* sequence as a basis for the possible amplification of a trypsin-like protease in *C. imicola*.

This part of the study had three main objectives. The first objective was to isolate the total RNA of about 500 *C. imicola* midges and synthesise cDNA. Secondly, to use the primers designed from *C. sonorensis* and *C. nubeculosus* sequences to amplify *C. imicola*, and finally, to analyse the sequence of the amplified DNA.

A novel protein in *C. imicola* was identified from the cDNA of *C. imicola* using *C. sonorensis* designed primers. The open reading frame of the protein consisted of 830 nucleotides, and translated, 276 amino acid residues. Furthermore, this protein was high in serine, proline and lysine residues resulting in a molecular weight of 29,7 kDa. The size and molecular weight of the protein was similar to *CsLTRYP3A* and therefore, a sequence alignment between the two sequences was performed. Sequence alignment has shown that the sequences were only 34.19% similar. Sequence homologies were determined using the BLAST software on the NCBI website. No significant similarities were found indicating that there are currently no proteins in the database with an identity to the *C. sonorensis* salivary protease. The protein identified is also not a protease since the online protease identifier software could not identify any protease domains in the sequence.

The conclusion of the work performed in this chapter is that *C. imicola* contains a protein similar in size and weight to the trypsin-like serine protease of *C. sonorensis.* Protease identifying software has confirmed that this protein is not a protease and sequence alignment with *CsLTRYP3A* has shown only 34.19% identity between the sequences. Since the genome of *C. imicola* is not yet sequenced, it is difficult to determine the type of protein sequenced and its underlying properties. In future studies of this protein, the sequence obtained can be used to design sequence specific primers containing restriction sites. This would enable the cloning of this protein into an expression vector, aiding in further investigations of this specific protein's functions and characteristics.

CHAPTER 6

Concluding summary and future prospects

6.1 Concluding summary

The Orbiviruses, AHSV and BTV, are primarily transmitted amongst their hosts through certain haematophagous midge vectors (*Culicoides spp.*) (Erasmus, 1973). Darpel *et al* (2011) identified a trypsin-like serine protease in the saliva of *C. sonorensis*, which plays a significant role in cleaving VP2 of BTV. It was also established that the protease activity of these midge species is capable of directly modifying the Orbivirus structure by cleaving VP2 (Marchi *et al.,* 1995; Van Dijk & Huismans, 1982). The selective cleavage of VP2 by trypsin-like serine proteases (Marchi *et al*., 1995) results in the generation of subsequent ISVPs (Van Dijk & Huismanns, 1982; Marchi *et al.,* 1995). This cleavage made a significant contribution to the infection ability of the virus (Darpel *et al.*, 2011). Manole *et al* (2012) recently identified a naturally occurring AHSV7 strain with a truncated VP2. Upon further investigation, this strain has also shown to be more infective than the common AHSV4 strain, since it outgrew AHSV4 in culture (Manole *et al*., 2012). Through the cleavage of these viral particles, the ability of adult *Culicoides* to transmit the virus is highly increased (Dimmocke, 1982; Darpel *et al*., 2011) and therefore, it is greatly significant to investigate the factors that influence the capability of arthropod-borne viruses to infect their insect vectors.

In this study, we argued that one of the vectors for AHSV, *C. imicola*, had a protease similar to the trypsin-like serine protease identified by Darpel *et al* (2011) in *C. sonorensis*. Furthermore, the *C. sonorensis* protease was recombinantly expressed and following incubation with AHSV4, VP2 of AHSV4 was cleaved. This cleavage could possibly create a virion similar to the truncated AHSV7 identified by Manole *et al* (2012). By investigating these factors, the viral replication process could possibly be better understood, aiding in the knowledge regarding the vectors' contribution to increased infectivity.

In order to further investigate the feasibility of the abovementioned factors, this study had four specific aims. This first objective this study was to detect a 29 kDa protease in the total protein extract of *C. imicola* that showed similar characteristics to the 29 kDa salivary protein identified in *C. sonorensis*. The total proteins of *C. imicola* were isolated from around 1200 midges and SDS-PAGE identified four major proteins of around 50 kDa, 40 kDa, 29 kDa and 25 kDa. Three minor proteins ranging between 95 kDa and 37 kDa were also identified. The 29 kDa protein identified in *C. imicola* showed proteolytic activity similar to the 29 kDa protein identified by Darpel *et al* (2011). Since the 29 kDa protein in *C. sonorensis* was identified to be a trypsin-like serine protease, the possibility of the simmilarly sized protein in *C. imicola* being a protease had to be investigated. Zymography confirmed that the detected 29 kDa protein in *C. imicola* had proteolytic activity and through inhibition with the trypsin inhibitor, antipain, this protease was confirmed to have trypsin-like characteristics.

The second objective was to determine if the trypsin-like serine protease of *C. sonorensis* could be recombinantly expressed as a proteolytically active protease. A recombinant *CulsonLTRYP* with a C-terminal histidine tag was generated using the cold expression vector, pColdIII. The histidine tagged *CulsonLTRYP* was expressed as a soluble protein through optimising the concentration of the inducer and the duration of expression. Proteolytic activity of the recombinantly expressed protease was confirmed on a gelatin-based SDS-PAGE, after which the protease was partially purified using metal chelate affinity chromatography. An unknown cellular protein was co-eluted with the protease but did not seem to interfere with the proteolytic activity of *CulsonLTRYP*. However, this still needs to be confirmed with the use of an optimised expression and purification system. Finally, the proteolytic character of the recombinantly expressed protease was established as trypsin-like through the use of a variety of protease inhibitors and substrates. Recombinant expression of the proteins responsible for cleaving VP2 of AHSV and BTV may aid in better understanding how the expression of specific genes contribute to the virus infection process, aiding in the future design of control measures for Arbovirus diseases.

The third objective was to determine if the partially purified recombinantly expressed protease is capable of cleaving AHSV VP2. As stated earlier, previous studies on Orbiviruses have demonstrated proteolytic cleavage of the outer capsid protein, VP2, through the treatment with trypsin, chymotrypsin or salivary proteases (Mertens *et al.,* 1987; Burroughs *et al.,* 1994; Darpel *et al*., 2011). Since the expressed recombinant protease was confirmed to be trypsin-like, it was decided to incubate the partially purified AHSV in the presence of the partially purified recombinant protease. A second passage strain of AHSV4 was cultivated in BHK-21 cells and purified. The purified virus was incubated with the recombinant protease and possible cleavage of VP2 was visualised using one-dimensional electrophoresis. It was expected that the exposure of the intact AHSV4 particles to the trypsin-like proteases would result in the cleavage of VP2, generating two or more cleavage products that remain associated with the outer capsid layer of the particle (Manole *et al*., 2012, Darpel *et al*., 2011). Unfortunately, as a result of low virus yield and the impurities present in the partially purified recombinant protease, cleavage of VP2 could not be determined. This resulted in the effects of recombinantly expressed proteases on VP2 of AHSV left undetermined.

The fourth and final objective of this study was to attempt to determine the sequence of the 29 kDa *C. imicola* protease detected earlier in the study by means of PCR. Complementary DNA was synthesised from isolated total RNA of *C. imicola*. Primers were designed based on sequence identities to the 29 kDa late trypsin-like serine protease of *C. sonorensis* on the NCBI database. These primers, non-specific to *C. imicola*, were then used in an attempt to amplify the 29 kDa protease in *C. imicola*. The trypsin-like serine protease in *C. imicola* consisted of around 830 nucleotides (Darpel *et al*., 2011). Thus, the expected size of a possible amplicon was to be in the region of 830 bp. Opitimisation of the PCR conditions resulted in non-specific amplification of three DNA segments. One of the *C. imicola* amplicons had the expected size of around 830 bp. Cloning of this amplicon into a TA-cloning vector enabled the generation of a large population of this amplified segment for sequence determination. Sequence analysis identified that the amplicon had an open reading frame of 830 bp. This corresponded to the 828 bp open reading frame of the reference sequence, *CsLTRYP3A* (Darpel *et al*.,

2011). The NCBI BLAST tool yielded no sequence identities to the newly discovered protein. The use of a web-based protease identifier tool confirmed that the amino acid sequence of the identified sequence did not belong to a protease family as it lacked any protease signature sequences such as those of serine proteases. To fully understand the nature of the protein identified as well as its possible role in virus transmission, further studies have to be performed on this protein.

The final conclusion for this study is that *C. imicola* do in fact contain a 29 kDa protease with trypsin-like properties similar to the 29 kDa protease identified by Darpel *et al* (2011). It is also possible to actively express the trypsin-like salivary protease of *C. sonorensis* in bacteria using a cold expression system. The expressed protein was confirmed to be trypsin-like with the use of different protease substrates and inhibitors. This expression profile has shown similarities to the activity profile of the 29 kDa protease of *C. imicola*. Therefore, it can be postulated that both the *Culicoides* vectors for AHSV and BTV, *C. sonorensis* and *C. imicola,* have a similar protease responsible for cleaving VP2. The data obtained from incubating the recombinant protease with purified virus was inconclusive as a low virus yield and partially pure recombinant protease impaired the visualisation of any possible protein digestion. However, the possibility still exists that the recombinant protease will cause digestion of VP2. However, further optimised virus recombinant protease purification is needed for confirmation. A 29 kDa protein present in *C. imicola* was also sequenced, and even though this sequence had a lot of similarities to the trypsin-like serine protease of *C. sonorensis*, no known sequences are currently linked to this protein. However, the possibility of it being a protease was ruled out since the sequence did not contain any protease domains.

Based on the work done in the study, we can conclude that there are proteases present in *C. imicola* that show significant similarities to the 29 kDa salivary protease of *C. sonorensis*. These similarities might suggest that these two *Culicoides* species are implementing closely related mechanisms in the cleavage of AHSV VP2, ultimately leading to an increase in virus infectivity and transmission.

6.2 Future prospects

Marchi and co-workers (1995) showed that the initial cleavage of AHSV VP2, does not in itself alter the infectivity of the virus, and can therefore be seen as only one of the factors involved in determining the infectivity levels for the insect host. Furthermore, it is speculated that additional processing or cleavage of VP2 is needed to increase the specificity of infectivity (Marchi *et al*., 1995). It is not yet known how important the modification of the virus structure by means of salivary proteins is in the transmission of the AHSV, since a variety of arthropods express protease inhibitors to facilitate feeding and for protection against pathogens. These protease inhibitors have shown to maximise the volume of blood secreted when feeding while preventing coagulation. These protease inhibitors been identified in the saliva of the vectors for BTV and AHSV, *C. sonorensis* and *C. nebucolosus* (Campbell *et al*., 2005; Langer *et al.*, 2007). It is thus important to determine if these inhibitors only partially inhibit the proteolytic enzymes responsible for the cleavage of VP2. Furthermore, Darpel *et al* (2011) speculated that these salivary inhibitors only partially inhibit the proteases responsible for VP2 cleavage and that full activity of the proteases would only commence in an ideal circumstance such as ambient temperatures. Future studies relating to these circumstances may aid in better understanding the specific mechanisms by which VP2 are cleaved and its contribution to an increase in infectivity and the transmission cycle of AHSV.

Possible future prospects might include:

1. Completing the purification of the recombinant *C. sonorensis* protease expressed in bacteria using C-terminal fused polyhistidine tags, followed by molecular exclusion chromatography to separate the purified proteins.
2. Determining the nucleic and amino acid sequences of the *C. imicola* and *C. bolitinos* homologues of the 29 kDa trypsin-like serine protease present in the saliva of *C. sonorensis*. The approach of sequence identification could include the extraction of the protein from a gel and sequencing of the proteolytically cleaved VP2 fragments using mass spectrometry.

3. Further analyses of AHSV vector salivary proteases, which trigger an increase in infectivity. This could include salivary protease digestion of all nine AHSV serotypes, comparing recombinant salivary protein digestion patterns by focusing on the digestion of VP2 that will generate intermediate sub-viral particles. The biological effect of the protein digestion on the viral infectivity of AHSV could subsequently also be investigated.
4. Identifying the exact protease cleavage sites on VP2 of all nine AHSV serotypes and compare the protease digested VP2s of AHSV serotypes 1-9 as well as the truncated VP2 of the naturally mutated AHSV7 strain.
5. Further characterisation of the biological effect and relevance of protease cleavage of AHSV VP2. This could be approached with the use of an AHSV reverse genetics system. A specific AHSV4/7 model with mutations identical to the proteolytically cleaved VP2 could be generated and the effect of the cleaved VP2 on replication and infectivity determined.
6. Comparing pathogenicity and infectivity of an AHSV VP2 mutant. The role of VP2 in pathogenesis and tissue tropism can be investigated using a chicken embryo model. The chicken embryo model will facilitate in determining if VP2 mutations have an effect on the pathogenicity of the virus.

REFERENCES

Alexander, K.A., Kat, P.W., House, J., O'Brien, S.J., Laurenson, M.K., McNutt, J.W. & Osburn B.I. 1995. African horse sickness and African carnivores. *Veterinary Microbiology,* 47:133-140.

Arnold, M., Patton, J.T. & McDonald, S.M. 2009. Culturing, storage, and quantification of rotaviruses. *Current protocols in microbiology*, 15C-3.

Andrew, M., Whiteley, P., Janardhana, V., Lobato, Z., Gould, A. & Coupar, B. 1995. Antigen specificity of the ovine cytotoxic T lymphocyte response to bluetongue virus. *Veterinary Immunology and Immunopathology,* 47:311-322.

Babé, L.M. & Craik, C.S. 1997. Viral proteases: Evolution of diverse structural motifs to optimise function. *Cell*, 91:427-430.

Barrett, A.J., Kembhavi, A.A., Brown, M.A., Kirschke, H., Knight, C.G., Tamai, M. & Hanada, K. 1982. L-trans-Epoxysuccinyl-leucylamido (4-guanidino) butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *Biochemistry Journal*, 201:189-198.

Bessette, P.H., Åslund, F., Beckwith, J. & Georgiou, G. 1999. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proceedings of the National Academy of Sciences*, 96(24):13703-13708.

Birnboim, H.C. & Doly, J. 1979. A rapid alkaline extraction procedure for screening of recombinant plasmid DNA. *Nucleic acid research,* 7:1513-1523.

Blackburn, N.K., Searle, L. & Phelps, R.J. 1985. Viruses isolated from *Culicoides* (Dipt. Ceratopognidae) caught at the veterinary research farm Mazowe, Zimbabwe. *Journal of Entomology Society South Africa*, 48:331-336.

Blanton, F.S., & Wirth, W.W. 1979. Arthropods of Florida and neighbouring land areas. Volume 10. The sand flies (*Culicoides*) of Florida (Diptera: Ceratopogonidae). *Arthropods of Florida and neighbouring land areas. Volume 10. The sand flies (Culicoides) of Florida (Diptera: Ceratopogonidae).*

Blum, H., Beier, H. & Gross, H. J. 1987. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis*, 8(2):93-99.

Boorman, J. 1974. The maintenance of laboratory colonies of *Culicoides variipennis* (Coq.), *C. nubeculosus* (Mg.) and *C. riethi Kieff*. (Diptera, Ceratopogonidae). *Bulletin of Entomological Research*, 64(03):371-377.

Boorman, J. 1993. Biting midges (Ceratopognidae). (*In* L. Papp. & B. Darvas., *eds.* Contributions to a manual of Palaeartctic diptera. Budapest. p. 349-368).

Boorman, J.P. & Wilkinson, P.J. 1983. Potential vectors of bluetongue in Lesbos, Greece. *Veterinary Record*, 113(17):395-396.

Both, G.W., Bellamy, A.R. & Mitchell, D.B. 1994. Rotavirus protein structure and function. *Current Topics in Microbiological lmmunology,* 185:67-105.

Bouayoune, H., Touti, J., El Hasnaoui, H., Baylis, M. & Mellor, P.S. 1998. The *Culicoides* vectors of African horsesickness virus in Morocco: distribution and epidemiological implications. (*In* Mellor, P.S., *et al*., *eds*. African Horse Sickness. Vienna. p. 113-125).

Bremer, C.W. 1976. A gel electrophoretic study of the protein and nucleic acid components of African horsesickness virus. *Onderstepoort Journal of Veterinary Research,* 43:193-199.

Bremer, C.W., Huismans, H. & Van Dijk, A.A. 1990. Characterisation and cloning of the African horsesickness virus genome. *Journal of general Virology*, 71:793-799.

Burrage, T.G. & Laegreid, W.W. 1994. African horsesickness: pathogenesis and immunity. *Comparative immunology, microbiology and infectious diseases*, 17(3):275-285.

Burroughs, J.N., O'Hara, R.S., Smale, C.J., Hamblin, C., Walton, A., Armstrong, R. & Mertens, P.P.C. 1994. Purification and properties of virus particles, infectious subviral particles, cores and VP7 crystals of African horsesickness virus serotype 9. *Journal of General Virology*, 75:1849-1857.

Campbell, C.L., Vandyke, K.A., Letchworth, G.J., Drolet, B.S., Hanekamp, T. & Wilson, W.C. 2005. Midgut and salivary gland transcriptomes of the arbovirus vector *Culicoides sonorensis* (Diptera: Ceratopogonidae). *Insect molecular biology*, 14(2):121-136.

Carpenter, F.H. 1967. Treatment of trypsin with TPCK. *Methods in enzymology*, 11(26):237.

Carr, S.M. & Griffiths, O.M. 1987. Preparative density-gradient ultracentrifigation of DNA. *Biochemistry Genetics*, 25:385-390.

Castillo-Olivares, J., Calvo-Pinilla, E., Casanova, I., Bachanek-Bankowska, K., Chiam, R., Maan, S. & Mertens, P.P.C. 2011. A modified vaccinia Ankara virus (MVA) vaccine expressing African horse sickness virus (AHSV) VP2 protects against AHSV challenge in an IFNAR−/− mouse model. *PloS one*, 6(1):16503.

Coetzer, J.A.W. & Erasmus, B.J. 1994. African horse sickness. *Infectious diseases of livestock with special reference to southern Africa*, 1:460-475.

Creighton, T.E., Darby, N.J. & Kemmink, J. 1996. The roles of partly folded intermediates in protein folding. *The FASEB journal*, 10(1):110-118.

Cronk, J.D. 2012. Protease: Biochemistry dictionary. www.guweb2.gonzaga.edu/faculty/cronk/biochem/ Date of access: 12 Aug. 2013.

Dardiri, A.H. & Salama, S.A. 1988. African horse sickness: an overview. *Journal of Equine Veterinary Science*, 8(1):46-49.

Darpel, K.E., Langner, K.F., Nimtz, M., Anthony, S.J., Brownlie, J., Takamatsu, H.H. & Mertens, P.P.C. 2011. Saliva proteins of vector *Culicoides* modify structure and infectivity of bluetongue virus particles. *PloS one*, 6(3):17545.

Derham, W. 1713. Physico theology, or a demonstration of the being and attributes of God from his works of creation. *London, UK: W. Innys*.

Diaz, M.R. & Panos Marti, M.P. 1968. La peste equina. *Bulletin - Office international des épizooties,* 70:647-662.

Dimmock, N.J. 1982. Initial stages in infection with animal viruses. *Journal of General Virology*, 59(1):1-22.

Diprose, J.M., Burroughs, J.N., Sutton, G.C., Goldsmith, A., Gouet, P., Malby, R. & Grimes, J.M. 2001. Translocation portals for the substrates and products of a viral transcription complex: the bluetongue virus core. *The EMBO journal*, 20(24):7229-7239.

Du Plessis, M. & Nel, L.H. 1997. Comparative sequence analysis and expression of the M6 gene, encoding the outer capsid protein VP5, of African horsesickness virus serotype nine. *Virus research*, 47(1):41-49.

Du Toit, R.M. 1944. The transmission of bluetongue and horsesickness by *Culicoides*. *Onderstepoort Journal of veterinary Science and animal Industry*, 19:7-16.

Dyce, A.L. 1969. The recognition of nulliparous and parous *Culicoides* (Diptera: Ceratopogonidae) without dissection. *Australian Journal of Entomology*, 8(1):11-15.

Erasmus, B.J. 1964. Some observation on the propagation of horse sickness virus in tissue culture. *Bulletin Office International des Epizooties*, 62:923-928.

Erasmus, B.J. 1973. Pathogenesis of African horsesickness. *In* Proceedings International Conference on Equine Infectious Diseases (Vol. 1973).

Forzan, M., Marsh, M. & Roy, P. 2007. Bluetongue virus entry into cells. *Journal of virology*, 81(9):4819-4827.

Fox, C.H. 1977. Staining cultures of cells. *Methods in Cell Science*, 3(2):583-585.

Freshney, R. 1987. Culture of Animal Cells: A Manual of Basic Technique. New York. Alan R. Liss Inc.

Gillespie, R.D., Mbow, M.L. & Titus R.G. 2000. The immunomodulatory factors of blood feeding arthropod saliva. *Parasite Immunology*, 22:319-331.

Gorman, B.M. & Taylor, J. 1985. Orbiviruses. (*In* Fields, B.N., *eds.,* Virology. New York: Raven Press. p. 907-925).

Gould, A.R. & Hyatt, A.D. 1994. The Orbivirus genus. Diversity, structure, replication and phylogenetic relationships. *Comparative immunology, microbiology and infectious diseases*, 17(3):163-188.

Graumann, K. & Premstaller, A. 2006. Manufacturing of recombinant therapeutic proteins in microbial systems. *Biotechnology journal*, 1(2):164-186.

Grimes, J., Basak, A., Roy, P. & Stuart, D. 1995. The crystal structure of bluetongue virus VP7: implications for virus assembly. *Nature,* 373:167-170.

Grimes, J.M., Burroughs, J.N., Gouet, P., Diprose, J.M., Malby, R., Ziéntara, S. & Stuart, D.I. 1998. The atomic structure of the bluetongue virus core. *Nature*, 395(6701):470-478.

Grubman, M.J., Appleton, J.A. & Letchworth, G.J. 1983. Identification of bluetongue virus type 17 genome segments coding for polypeptides associated with virus neutralisation and intergroup reactivity. *Virology,* 131:355-366.

Grubman, M.J. & Lewis, S.A. 1992. Identification and characterization of the structural and nonstructural proteins of African horsesickness virus and determination of the genome coding assignments. *Virology*, 186(2):444-451.

Henning, M.M. 1956. African Horsesickness, Perdesiekte, Pestis Equorum. (*In* Animal Diseases in South Africa. Central News Agency Ltd. p. 785-808).

Hess, W.R. 1988. African horse sickness. (*In* Monath, T.P., *ed.* Arboviruses, Epidemiology and Ecology. Florida. p. 1-18).

Heussen, C. & Dowdle, E.B. 1980. Electrophoretic analysis of plasminogen activators in polyacrylamide gels containing sodium dodecyl sulfate and copolymerized substrates. *Analytical biochemistry*, 102(1):196-202.

Hocking, B. 1971. Blood-sucking behavior of terrestrial arthropods. *Annual Review of Entomology*, 16(1):1-28.

House, J.A. 1993b. Recommendations for African horse sickness vaccines for use in non-endemic areas. *Revue d Elevage et de Medicine Veterinaire des pays Tropicaux.* 46:77-81.

Howell, P.G. 1960. A preliminary antigenic classification of strains of bluetongue virus. *Onderstepoort Journal of Veterinary Research,* 28:357-363.

Howell, P.G. 1962. The isolation and identification of further antigenic types of African horsesickness virus. *Onderstepoort Journal of Veterinary Research,* 29(2):139-149.

Howell, P.G. 1963. African horse sickness. *Emerging diseases of animals*, 61:71-108.

Huismans, H. & Verwoerd, D.W. 1973. Control of transcription during the expression of the bluetongue virus genome. *Virology*, 52(1):81-88.

Huismans, H. & Els, H.J. 1979. Characterization of the tubules associated with the replication of three different Orbiviruses. *Virology*, 92(2):397-406.

Huismans, H., van der Walt, N.T., Cloete, M. & Erasmus, B.J. 1987. Isolation of a capsid protein of bluetongue virus that induces a protective immune response in sheep. *Virology*, 157(1):172-179.

Huismans, H., van Dijk, A.A. & Bauskin, A.R. 1987a. In vitro phosphorylation and purification of a non-structural protein of bluetongue virus with affinity for single-stranded RNA. *Journal of Virology,* 61(11):3589-3595.

Huismans, H., Van Dijk, A.A. & Els, H.J. 1987b. Uncoating of parental bluetongue virus to core and subcore particles in infected L cells. *Virology,* 157:180-188.

Huismans, H. & Van Dijk, A.A. 1990. Bluetongue virus structural components. (*In* Roy, P. & Gorman, B.M., *eds.* Bluetongue Viruses. Berlin: Heidelberg. p. 21-41).

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.

Hutcheon, D. 1902. Malarial catarrhal fever of sheep. *Veterinary Record*, 14:629-633.

Hutchinson, I.R. 1999. The role of VP7 (T13) in initiation of infection by bluetongue virus. University of Hertfordshire. (Thesis - PhD).

Hyatt, A.D., Zhao, Y. & Roy, P. 1993. Release of bluetongue virus-like particles from insect cells is mediated by BTV non-structural protein NS3/NS3A. *Virology,* 193:592-603.

International Committee on Taxonomy of Viruses (ICTV). 2012. Virus taxonomy: 2012 release. http://www.ictvonline.org/virusTaxonomy.asp?version=2012 Date of access: 2 Oct. 2013.

Iwata, H., Chuma, T. & Roy, P. 1992a. Characterisation of the genes encoding two of the major capsid proteins of epizootic haemorrhagic disease virus indicates a close genetic relationship to bluetongue virus. *Journal of General Virology*, 173:915-24.

Jacobsson, K. 1955. I. Studies on the determination of fibrinogen in human blood plasma. II. Studies on the trypsin and plasmin inhibitors in human blood serum. *Scandinavian journal of clinical and laboratory investigation*, 7:3.

Kaname, Y., Celma, C.C., Kanai, Y. & Roy, P. 2013. Recovery of African horse sickness virus from synthetic RNA. *Journal of General Virology*, 94(10):2259-2265.

Kar, A.K., Ghosh, M. & Roy, P. 2004. Mapping the assembly pathway of Bluetongue virus scaffolding protein VP3. *Virology*, 324(2):387-399.

Katunuma, N. & Kominami, E. 1995. Structure, properties, mechanisms, and assays of cysteine protease inhibitors: Cystatins and E-64 derivatives. *Methods in enzymology*, 251:382-397.

Kiuchi, A., Rao, C.D. & Roy, P. 1983. Analysis of bluetongue viral RNA sequences. (*In* Compans, R.W. & Bishop, D.H.L., *eds*. Double- stranded RNA viruses. New York. p. 55-64).

Klenk, H.D. & Garten, W. 1994. Host cell proteases controlling virus pathogenicity. *Trends in microbiology*, 2(2):39-43.

Laegreid, W.W., Skowronek, A., Stone-Marschat, M. & Burrage, T. 1993. Characterisation of virulence variants of African horsesickness virus. *Virology,* 195:836-839.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259):680-685.

Langner, K.F., Darpel, K.E., Denison, E., Drolet, B.S., Leibold, W., Mellor, P.S. & Greiser-Wilke, I. 2007. Collection and analysis of salivary proteins from the biting midge *Culicoides nubeculosus* (Diptera: Ceratopogonidae). *Journal of medical entomology*, 44(2):238-248.

Lecatsas, G. 1968. Electron microscopic study of the formation of bluetongue virus. *Onderstepoort journal of veterinary research*, 35(1):139.

Loughran, S.T., Loughran, N.B., Ryan, B.J., D'Souza, B.N. & Walls, D. 2006. Modified His- tag fusion vector for enhanced protein purification by immobilized metal affinity chromatography. *Analytical Biochemistry*, 355:148-150.

MacLachlan, N.J. & Guthrie, A.J. 2010. Re-emergence of bluetongue, African horse sickness, and other Orbivirus diseases. *Veterinary research*, 41(6):35.

Manole, V., Laurinmäki, P., van Wyngaardt, W., Potgieter, C.A., Wright, I.M., Venter, G. J., van Dijk, A.A., Sewell, B.T. & Butcher, S.J. 2012. Structural insight into African horsesickness virus infection. *Journal of virology*, 86(15):7858-7866.

Marchi, P.R., Rawlings, P., Burroughs, J.N., Wellby, M., Mertens, P.P.C., Mellor, P.S. & Wade-Evans, A.M. 1995. Proteolytic cleavage of VP2, an outer capsid protein of African horse sickness virus, by species-specific serum proteases enhances infectivity in *Culicoides*. *Journal of general virology*, 76(10):2607-2611.

Marciniszyn, J., Hartsuck, J. & Tang, J.J.N.J. 1976. Mode of inhibition of acid proteases by pepstatin. *Journal of Biological Chemistry*, *251*(22):7088-7094.

Marchuk, D., Drumm, M., Saulino, A. & Collins, F.S. 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Research*, 19(5):1154-1154.

Maree, F.F. & Huismans, H. 1997. Characterization of tubular structures composed of nonstructural protein NS1 of African horsesickness virus expressed in insect cells. *Journal of general virology*, 78(5):1077-1082.

Maree, S., Durbach, S., Maree, F.F., Vreede, F. & Huismans, H. 1998b. Expression of the major core structural proteins VP3 and VP7 of African horse sickness virus, and production of core-like particles. *Archives of Virology Supplement,* 14:203-209.

Martinez-Torrecuadrada, J. L., Iwata, H., Venteo, A., Casal, I. & Roy, P. 1994. Expression and characterization of the two outer capsid proteins of African horsesickness virus: the role of VP2 in virus neutralization. *Virology*, 202(1):348-359.

Martínez-Torrecuadrada, J.L. & Casal, J.I. 1995. Identification of a linear neutralization domain in the protein VP2 of African horse sickness virus. *Virology*, 210(2):391-399.

Martínez-Torrecuadrada, J.L., Díaz-Laviada, M., Roy, P., Sánchez, C., Vela, C., Sánchez-Vizcaíno, J.M. & Casal, J.I. 1996. Full protection against African horsesickness (AHS) in horses induced by baculovirus-derived AHS virus serotype 4 VP2, VP5 and VP7. *Journal of general virology*, 77(6):1211-1221.

Matsuo, E. & Roy, P. 2009. Bluetongue virus VP6 acts early in the replication cycle and can form the basis of chimeric virus formation. *Journal of virology*, 83(17):8842-8848.

Mathieu, B. 2010. What is African horsesickness?. http://www.sanidadanimal.info/sanidadanimal/en/actividades/emerging-online/african-horse-sickness/162-que-es-la-peste-equina-africana.html Date of access: 16 Jul. 2013.

McIntosh, B.M. 1958. Immunological types of horsesickness virus and their significance in immunization. *Onderstepoort Journal of Veterinary Research*, 27:465-538.

Meiswinkel, R., Nevill E. & Venter G. 1994. Vectors: *Culicoides* spp. (*In* Coetzer J., Thomson, G. & Tustin R., *eds.* Infectious diseases of livestock with special reference to Southern Africa. Cape Town, p. 68–89).

Mellor, P.S. 1994. Epizootiology and vectors of African horse sickness virus. *Comparative Immunology, Microbiology and Infectious Diseases*, 17:287-296.

Mellor, P.S., Boorman, J. & Jennings, M. 1975. The multiplication of African horsesickness virus in two species of *Culicoides* (Diptera, Ceratopogonidae). *Archives of virology*, 47(4):351-356.

Mellor, P.S., Osborne, R. & Jennings, D.M. 1984. Isolation of bluetongue and related viruses from *Culicoides* species in the Sudan. *Journal of Hygiene*, 93:621-628.

Mellor, P.S., Walton, T.E. & Osburn, B.I. 1992. *Culicoides* as potential Orbivirus vectors in Europe. (*In* Bluetongue, African horse sickness, and related Orbiviruses: Proceedings of the Second International Symposium. CRC Press, Inc. p. 278-283).

Mellor, P.S., Boorman, J. & Baylis, M. 2000. *Culicoides* biting midges: their role as arbovirus vectors. *Annual review of entomology*, 45(1):307-340.

Mellor, P.S. & Mertens, P.P.C. 2008. African horse sickness viruses. *Encyclopedia of Virology*, 2008:37-43.

Mertens, P.P.C. 2004. The dsRNA viruses. *Virus Research,* 101(1):3-13.

Mertens, P.P.C., Brown, F. & Sangar, D.V. 1984. Assignment of the genome segments of bluetongue virus type 1 to the proteins they encode. *Virology,* 135:207-217.

Mertens, P.P.C., Burroughs, J.N. & Anderson, J. 1987. Purification and Properties of Virus-Particles, Infectious Subviral Particles, and Cores of Bluetongue Virus Serotype-1 and Serotype-4. *Virology,* 157:375-386.

Mertens, P.P.C., Pedley, S., Cowley, J., Burroughs, J.N., Corteyn, A.H., Jeggo, M.H., Jennings, D.M. & Gorman, B.M. 1989. Analysis of the roles of bluetongue virus outer capsid proteins VP2 and VP5 in determination of virus serotype. *Virology,* 170:561-565.

Mertens, P.P.C., Burroughs, J.N., Walton, A., Wellby, M.P., Fu, H., O'Hara, R.S., & Mellor, P.S. 1996. Enhanced Infectivity of Modified Bluetongue Virus Particles for Two Insect Cell Lines and for two *Culicoides* Vector Species. *Virology*, 217(2):582-593.

Mertens P.P.C., Maan S., Samuel A. & Attoui H. 2005. Orbivirus*, Reoviridae*. (*In* Fauquet C.M., Mayo M.A., Maniloff J., Desselberger U. & Ball L.A., *eds*. Virus Taxonomy, VIIIth Report of the ICTV. London. p. 466-483).

Mertens, P.P.C & Diprose, J. 2004. The bluetongue virus core: a nano-scale transcription machine. *Virus research*, 101(1):29-43.

Mirchamsy, H. & Rapp, F. 1968. A new overlay for plaquing animal viruses. *Proceedings of the Society for Experimental Biology and Medicine,* 129(1):13-17.

Mizukoshi, N., Sakamoto, K., Iwata, A., Tsuchiya, T., Ueda, S., Apiwatnakorn, B., Kamada, M. & Fukusho, A. 1993. The complete nucleotide sequence of African horsesickness virus serotype 4 (vaccine strain) segment 4, which encodes the minor core protein VP4. *Virus Research,* 28:299-306.

Morgan, J.F., Morton, H.J. & Parker, R.C. 1950. Nutrition of Animal Cells in Tissue Culture. I. Initial Studies on a Synthetic Medium. *Proceedings of the Society for Experimental Biology and Medicine,* 73(1):1-8.

Mouie, L. 1896. Historie de la medicine vétérinaire. *Maulde*, Paris. p. 38.

Mullins, J. 2010. Use of expression systems in research and the pharmaceutical industry. http://www.rothamsted.ac.uk/notebook/phar.htm Date of access: 11 Nov. 2013.

Neumann, E., Schaefer-Ridder, M., Wang, Y. & Hofschneider, P.H. 1982. Gene transfer into mouse lyoma cells by electroporation in high electric fields. *The EMBO journal*, 1(7):841.

Nevill, E.M. & Anderson, D. 1972. Host preferences of *Culicoides* midges (Diptera: Ceratopogonidae) in South Africa as determined by precipitin tests and light trap catches. *The Onderstepoort journal of veterinary research*, 39(3):147.

Nevill, E.M., Venter, G.J. & Edwardes, M. 1992. Potential *Culicoides* vectors of livestock Orbiviruses in South Africa. (*In* Bluetongue, African horse sickness, and related Orbiviruses: Proceedings of the Second International Symposium. p. 306-313).

Nibert, M.L. 1998. Structure of Mammalian *Orthoreovirus* Particles. *Current Topics in Microbiology and Immunology,* 233:1-32.

Nibert, M.L., Schiff, L.A. & Fields, B.N. 2001. Reoviruses and their replication. (*In* Fields virology, 4th ed. Philadelphia, Pa: Lippincott Williams & Wilkins, p. 1679-1728).

O'Hara, R.S., Meyer, A.J., Burroughs, J.N., Pullen, L., Martin, L.A. & Mertens, P.P.C. 1998. Developments of a mouse model system, coding assignments and identification of the genome segments controlling virulence of African Horse sickness virus serotypes 3 and 8. *Archives of Viral Supplement*, 14:259-279.

Oellermann, R.A., Els, H.J. & Erasmus, B.J. 1970. Characterisation of African horsesickness virus. *Archiv für die Gesamte Virusforsch,* 29:163-74.

Oehler, S., Eismann, E.R., Krämer, H. & Müller-Hill, B. 1990. The three operators of the lac operon cooperate in repression. *The EMBO journal*, *9*(4):973.

OIE (2013) Classification of Diseases. *Bulletin - Office international des épizooties*. World organisation for Animal Health. *http://www.oie.int.*

Owen, N.C. & Mum, E.K. 1966. Observations on a strain bluetongue virus by electron microscopy. *Onderstepoort Journal of veterinary research,* 33:9-14.

Palmer, T. 2001. Enzymes: Biochemistry, Biotechnology and Clinical Chemistry. Chichester: Horwood Publishing Limited.

Paton, T. 1863. The ‘‘horse sickness’’ of the Cape of Good Hope. *Veterinarian*, 36:489-494.

Pedley, S., Mohamed, M.E.H. & Mertens, P.P.C. 1988. Analysis of genome segments from six different isolates of bluetongue virus using RNA-RNA hybridisation: a generalised coding assignment for bluetongue viruses. *Virus Research,* 10:381-390.

Pitchford, N. & Theiler, A. 1903. ‘Investigations into the nature and causes of horsesickness’. *Agricultural Journal of the Cape of Good Hope,* 23:153-156.

Potgieter, A.C. 2004. Cloning viral dsRNA genomes: analysis and application. Potchefstroom: NWU. (Thesis - PhD).

Potgieter, A.C., Page, N.A., Liebenberg, J., Wright, I.M., Landt, O. & van Dijk, A.A. 2009. Improved strategies for sequence-independent amplification and sequencing of viral double-stranded RNA genomes. *Journal of General Virology*, 90(6):1423-1432.

Prasad, B.V.V., Yamaguchi, S. & Roy, P. 1992. Three dimensional structure of single shelled BTV. *Journal of Virology,* 66:*2135-2142.*

Purification of His-tag proteins: User manual. 2011. *Protino® Ni-TED packed columns*. Macherey-Nagel, Germany. Rev.04.

Rafyi, A. 1961. Horsesickness. *Bulletin - Office international des épizooties*, 56:216-250.

Rahman, S., Ahmad, A. & Bhatti, S. 2006. Comparative growth promoting efficacy of fetal calf serum (FCS) for baby hamster kidney-21 (BHK-21) cell line. *Journal of Animal and Plant Science*, 16:1-2

Ramadevi, N., Burroughs, N.J., Mertens, P.P.C., Jones, I.M. & Roy, P. 1998. Capping and methylation of mRNA by purified recombinant VP4 protein of bluetongue virus. *Proceedings of the National Academy of Sciences*, 95(23):13537-13542.

Rao, C.D., Kiuchi, A. & Roy, P. 1983. Homologous terminal sequences of the genome double- stranded RNAs of bluetongue virus. *Journal of Virology,* 46:378-383.

Rawlings, N.D. & Barrett, A.J. 1994. Families of serine peptidases. *Methods in Enzymology,* 244:19-61.

Rawlings, P., Pro, M.J., Penal Onega, M.D. & Cpaela, R. 1997. Spatial and seasonal distribution of *Culicoides imicola* in Iberia in relation to the transmission of African horse sickness virus. *Medical and Veterinary Entomology*, 11:49-57.

Rawlings, N., Tolle, D. & Barrett, A. 2004. Evolutionary families of peptidase inhibitors. *Biochemistry Journal*, 378:705-716.

Reed, L.J. & Muench, H. 1938. A simple method of estimating fifty per cent endpoints. *American Journal of Epidemiology*, 27(3):493-497.

Ribeiro, J.M. & Francischetti, I.M. 2003. Role of Arthropod Saliva in Blood Feeding: Sialome and Post-Sialome Perspectives. *Annual Review of Entomology*, 48(1):73-88.

Roux, K.H. 1995. Optimization and troubleshooting in PCR. *Genome Research*, *4*(5):185-194.

Roy, P. 1992. From genes to complex structures of bluetongue virus and their efficacy as vaccine. *Veterinary Microbiology,* 33:155-168.

Roy, P. 2001. Orbiviruses. (*In* Fields, B.N., Knipe, D.N., Howley, P.M. & Griffin, D.E., *eds*. Virology. Philadelphia. p. 1679-1728).

Roy, P. 2011. Orbivirus. (*In* Tidona, C.A. & Darai, G., *eds*. The Springer Index of Viruses. Springer, New York. p. 1603-1610).

Roy, P., Fukusho, A., Ritter, D.G. & Lyons, D. 1988. Evidence for genetic relationship between RNA and DNA viruses from the sequence homology of a putative polymerase gene of bluetongue virus with that of vaccinia virus: conservation of RNA polymerase genes from diverse species. *Nucleic Acids Research,* 16:11759-11767.

Roy, P., Mertens, P.P.C. & Casal, I. 1994b. African horse sickness virus structure. *Comparative Immunology, Microbiology and Infectious Diseases,* 17:243-73.

Roy, P., Bishop, D.H., Howard, S., Aitchison, H. & Erasmus, B. 1996. Recombinant baculovirus-synthesised African horsesickness virus (AHSV) outer-capsid protein VP2 provides protection against virulent AHSV challenge. *Journal of General Virology,* 77:2053-2057.

Roy, P., Mikhailov, M. & Bishop, D.H. 1997. Baculovirus multigene expression vectors and their use for understanding the assembly process of architecturally complex virus particles. *Gene,* 190(1):119-129.

Roy, P. & Sutton, G. 1998. New generation of African horsesickness virus vaccines based on structural and molecular studies of the virus particles. *Archives of Virology Supplement*, 14:177-202.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G. T. & Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239(4839):487-491.

Saluta, M. & Bell, P.A. 1998. Troubleshooting GST fusion protein expression in E. coli. *Life Science News*, 1:1-3.

Salvesen, G.S., & Nagase, H. 2001. Inhibition of proteolytic enzymes. (*In* R. Beynon & J.S. Bond., *eds.* Proteolytic Enzymes: a practical approach. Oxford: Oxford University Press. p. 340

Sambrook, J. & Russell, D.W. 2001. Molecular cloning : A laboratory manual. 3rd ed. Cold Spring Harbour Laboratory Press.

Seegers, L. 2010. African Horse Sickness Precautions. www.perseveranceendurancehorses.wordpress.com/2011/05/23/african-horse-sickness-precautions-november-2010 Date of access: 24 Sept. 2013.

Schatz, A., Bugle, E. & Waksman, S.A. 1944. Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria. *Proceedings of the Society for Experimental Biology and Medicine.* 55(1):66-69.

Schmidt, G.M., Krehbiel, J.D., Coley, S.C. & Leid, R.W. 1982. Equine onchocerciasis: lesions in the nuchal ligament of Midwestern US horses. *Veterinary Pathology Online*, 19(1):16-22.

Smooker, P.M., Jayaraj, R., Pike, R.N. & Spithill, T.W. 2010. Cathepsin B proteases of flukes: the key to facilitating parasite control?. *Trends in parasitology*, *26*(10):506-514.

Shi, X. Z., Ren, Q., Zhao, X. F., & Wang, J. X. 2009. Expression of four trypsin-like serine proteases from the Chinese shrimp, *Fenneropenaeus chinensis*, as regulated by pathogenic infection. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 153(1):54-60.

Stauber, N., Martinez-Costas, J., Sutton, G., Monastyrskaya, K. & Roy, P. 1997. Bluetongue virus VP6 protein binds ATP and exhibits an RNA-dependent ATPase function and a helicase activity that catalyze the unwinding of double-stranded RNA substrates. *Journal of Virology,* 171:7220-7226.

Stoker, M. & Macpherson, I. 1964. Syrian hamster fibroblast cell line BHK21 and its derivatives. *Nature,* 26(203):155-157.

Stark, M.J. 1987. Multicopy expression vectors carrying the *lac* represser gene for regulated high-level expression of genes in *Escherichia coli*. *Gene*, *51*(2):255-267.

Strober, W. 2001. Trypan blue exclusion test of cell viability. *Current protocols in immunology*, A-3B.

Szybalski, W. 1960. Sampling of virus particles and macromolecules sedimented in an equilibrium density gradient. *Experientia*, *16*(4):164-164.

Tang, J. & Wong, R.N. 1987. Evolution in the structure and function of aspartic proteases. *Journal of cellular biochemistry*, 33(1):53-63.

Terpe, K. 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied microbiology and biotechnology*, 60(5):523-533.

Terpe, K. 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied microbiology and biotechnology*, 72(2):211-222.

Theiler, A. 1901. Die Zuafrikanische Pferdesterbe. *Deutsche Tierarztliche Wochenschrift,* 9:201-203.

Theiler, A. 1921. African horse sickness. *Science Bulletin.* No 19. Department of Agriculture. South Africa. Pretoria.

Theron, J., Uitenweerde, J.M., Huismans, H. & Nel, L.H. 1994. Comparison of the expression and phosphorylation of the non-structural protein NS2 of three different Orbiviruses: evidence for the involvement of a ubiquitous cellular kinase. *Journal of General Virology,* 75:3401-11.

Theron, J. & Nel, L.H. 1997. Stable protein-RNA interaction involves the terminal domains of bluetongue virus mRNA, but not the terminally conserved sequences. *Virology,* 229:134-42.

Titus, R.G. & Ribeiro, J.M.C. 1990. The role of vector saliva in transmission of arthropod-borne disease. *Parasitology Today*, 6(5):157-160.

Umezawa, H. 1972. Pepstatin. *Enzyme inhibitors of microbial origin*, 1972:34-50.

Urakawa, T. & Roy, P. 1988. Bluetongue virus tubules made in insect cells by recombinant baculoviruses: expression of the NS1 gene of bluetongue virus serotype 10. *Journal of virology*, 62(11):3919-3927.

Urbano, P. & Urbano, F.G. 1994. The *Reoviridae* Family. *Comparative immunology, microbiology and infectious diseases,* 17:151-161.

Van Ark, H. & Meiswinkel, R. 1992. Subsampling of large light trap catches of *Culicoides* (Diptera: Ceratopogonidae). *The Onderstepoort journal of veterinary research*, 59(3):183.

Van der Walt, N.T. 1980. A haemagglutination and haemagglutination inhibition test for bluetongue virus. *Onderstepoort Journal of Veterinary Research*, 47(2):113-117.

Van Dijk, A.A. & Huismans, H. 1982. The effect of temperature on the *in vitro* transcriptase reaction of bluetongue virus, epizootic haemorrhagic disease virus and African horsesickness virus. *Onderstepoort Journal of Veterinary Research,* 49:227-232.

Van Dijk, A.A. & Huismans, H. 1988. *In vitro* transcription and translation of bluetongue virus mRNA. *Journal of General Virology*, 169:573-81.

Van Rensburg, L.B.J., De Clerk J., Groenewald H.B. & Botha W.S. 1981. An outbreak of African horse sickness in dogs. *Journal of South African Veterinary Association*, 52(1981):323-325.

Van Staden, V. & Huismans, H. 1991. A comparison of the genes which encode non-structural protein NS3 of different Orbiviruses. *The Journal of general virology*, 72:1073.

Venter, G.J., Labuschagne, K., Hermanides, K.G., Boikanyo, S.N.B., Majatladi, D.M. & Morey, L. 2009. Comparison of the efficiency of five suction light traps under field conditions in South Africa for the collection of *Culicoides* species. *Veterinary Parasitology*, 166(3):299-307.

Verwoerd, D.W. 1969. Purification and characterisation of bluetongue virus. *Virology,* 38: 203-212.

Verwoerd, D.W. & Erasmus, B.J. 1994. Bluetongue. (*In* Coetzer, J.A.W., Thomson, G.R. & Tustin, R.C., *eds*. Infectious diseases of livestock with special reference to Southern Africa. Cape Town: Oxford University Press. *p.* 443-459).

Verwoerd, D.W., Els, H.J., De Villiers, E.M. & Huismans, H. 1972. Structure of the bluetongue virus capsid. *Journal of Virology,* 10:783-794.

Vogelstein, B. & Gillespie, D. 1979. Preparative and analytical purification of DNA from agarose. *Proceedings of the National Academy of Sciences USA.* 76:615-619.

Voet, D. & Voet, J.G. 1995. Biochemisty. 2nd ed. New york, John Wiley & Sons.

Vreede, F.T. & Huismans, H. 1994. Cloning, characterisation and expression of the gene that encodes the major neutralisation-specific antigen of African horsesickness virus serotype 3. *Journal of General Virology, 175:3629-33.*

Vreede, F.T. & Huismans, H. 1998. Sequence analysis of the RNA polymerase gene of African horse sickness virus. *Archives of Virology*, 143:413-419.

Wade-Evans, A.M., Pullen, L., Hamblin,C., O'Hara,R. Burroughs, J.N. & Mertens, P.P.C. 1997. African horsesickness virus VP7 sub-unit vaccine protects mice against a lethal, heterologous serotype challenge. *Journal of General Virology,* 78:1611-1616.

Wagner, W., Möhrlen, F. & Schnetter, W. 2002. Characterization of the proteolytic enzymes in the midgut of the European Cockchafer, *Melolontha melolontha* (Coleoptera: Scarabaeidae). *Insect biochemistry and molecular biology*, *32*(7):803-814.

Walker, A.R. & Davies, F.G. 1971. A preliminary survey of the epidemiology of bluetongue in Kenya. *Journal of Hygiene*, 69(01):47-60.

Wan, M., Takagi, M., Loh, B. & Imanaka, T. 1995. Comparison of HIV-1 protease expression in different fusion forms. *Biochemistry and molecular biology international,* 36(2):411-419.

Wellby, M.P., Baylis, M., Rawlings, P. & Mellor, P.S. 1996. Effect of temperature on survival and rate of virogenesis of African horse sickness. *Bulletin of entomological research*, 86:715-720.

Wikel, S.K. 1999. Modulation of the host immune system by ectoparasitic arthropods. *BioScience*, 49(4):311-320.

Wilson, A., Mellor, P.S., Szmaragd, C. & Mertens, P.P.C. 2009. Adaptive strategies of African horsesickness virus to facilitate vector transmission. *Veterinary Research,* 40:16.

Zhang, X., Boyce, M., Bhattacharya, B., Zhang, X., Schein, S., Roy, P. & Zhou, Z.H. 2010. Bluetongue virus coat protein VP2 contains sialic acid-binding domains, and VP5 resembles enveloped virus fusion proteins. *Proceedings of the National Academy of Sciences*, 107(14):6292-6297.

APPENDIX I

Materials used in this study

Item	Product no.	Supplier
6x Orange Loading dye	R0631	Fermentas
Acrylamide	1.00209.100	Merck
Agar	BX.10.500	Merck
Agarose	#D1-LE	White-Sci
Amido black	0210056325	MP Biomedical
Ammonium persulfate	A3678	Sigma
Amphotericin B	15290-018	Life Technologies
Ampicillin	A9393	Sigma
AMV-RT	M5101	Promega
Antipain	11004646001	Roche
Benzonase	70746-3	Merck
Bio-Rad GenePulser electroporation cuvettes	165-2089	Bio-Rad
Boc-Asp(OBzl)-Pro-Arg-AMC	3139-v5mg	PeptaNova
BOC-Val-Pro-Arg-AMC	3093-v5mg	PeptaNova
Brij 35	ICN10111101	MP Biomedical
BugBuster™	70584-4	Novagen
BZ-Arg-AMC	3092-v5mg	PeptaNova
Caesium chloride	289329-25g	Sigma
Chloroform	288306-1L	Sigma
Chymostatin	11004638001	MP Biomedical
Coomassie brilliant blue R-250	B8647	Sigma
Disposable cuvettes	165-2089	Ratiolab
DMSO	472301	Sigma
DTT	43815	Sigma
DMEM	11965-118	Life Technologies
E-64	E3132-5mg	Sigma
EDTA	3658	Fluka
Ethanol 100%	100,983.25	Merck

Ethidium bromide	160539	Sigma
ExTaq DNA polymerase	RR001	Takara
Fermentas insTAclone PCR cloning kit	K1214	Fermentas
Fermentas pageruler #SM1173	#SM1173	Fermentas
Fermentas pageruler #SM26619	#SM1173	Fermentas
Fetal Bovine Serum	10099-141	Life Technologies
FluorNunc® 96-well fluorometry plates	P8741	Sigma
Formaldehyde 37%	1039991000	Merck
Gelatin	G1890-100g	Sigma
Glacial acetic acid	BB100017P	BDH
Glycerol	49780	Fluka
Hydrochloric acid	SAAR3063040LP	Merck
IPTG	V395A	Promega
Isopropanol	1.09634.2500	Merck
Leupeptin	11017101001	Roche
Lysozyme	71110-4	Novagen
Macherey-Nagel Nucleospin II PCR cleanup and gel extraction kit	740609.250	Macherey-Nagel
Macherey-Nagel Protino® Ni-TED 2000 kit	745120.25	Macherey-Nagel
Macherey-Nagel NucleoSpin® Extract II	740609.250	Macherey-Nagel
Macherey-Nagel Nucleospin® plasmid DNA Miniprep kit	740588.250	Macherey-Nagel
Magnesium chloride	8.14733.0500	Merck
Methanol	1.06009.2500	Merck
N,N' Bisacrylamide	130672	Sigma
NdeI	# ER 0581	Thermo Scientific
Non-essential Amino Acids	11140-035	Life Technologies
Non-reducing loading buffer	39001	Thermo Scientific
O'Gene Ruler DNA ladder mix	#R0491	Fermentas

Phosphate buffered saline	P3813-10PAK	Sigma
Penicillin and Streptomycin	15070-063	Life Technologies
Pepstatin A	P5318-5mg	Sigma
Potassium chloride	AB004936	Merck
Protein loading buffer	39000	Thermo Scientific
Protein molecular size marker	SM26620	Fermentas
RNase A (Solution or powder)	R 6,513	Sigma
RNase-free water	W4502-1L	Sigma
RNasin® ribonuclease inhibitor	N2511	Promega
SalI	#ER 0641	Thermo Scientific
SDS	L4390	Sigma
Silver nitrate	1015100050	Merck
Sodium azide	26628-22-8	Fisher Scientific
Sodium carbonate	451614-25g	Sigma
Sodium chloride	S9888	Sigma
Sodium hydroxide	1.06498.0500	Merck
Sodium thiosulfate petahydrate	LC24990	Labchem
Sucrose	S9378-500g	Sigma
T4 DNA ligase	EL 0011	Fermentas
T4 DNA ligase buffer (x10)	#3705	Fermentas
TEMED	1.10732-026	Merck
TPCK	T4376-100mg	Sigma
Tri-Reagent	RT111	MRC
Tris Base	11814273001	Roche
Triton X-100 solution	X100-500ml	Sigma
Trypan blue	T6146-25g	Sigma
Tryptose	1.10676	Merck
Vivaspin 6 ultrafiltration system	VS0691	Sartorius
Yeast extract	C68	Biolab
Z-Pyr-Gly-Arg-AMC	3138-v	PeptaNova

18Ω water was prepared on site for use in general applications