Recombinant cathepsin L-like cysteine proteases for species-specific diagnosis of animal African trypanosomiasis

by

Ziphezinhle Elaine Mbhele

BSc. (Hons) Biochemistry

Submitted in fulfilment of the academic requirements for the degree of

Master of Science in Biochemistry

Biochemistry, School of Life Sciences

University of KwaZulu-Natal

Pietermaritzburg

South Africa

2019

Preface

The experimental work described in this dissertation was carried out in the School of Life Sciences, University of KwaZulu-Natal, Pietermaritzburg, from January 2017 to July 2019, under the supervision of Professor THT Coetzer. The studies represent original work by the author and have not otherwise been submitted in any other form to another University. Where use has been made of the work of others, it has been duly acknowledged in the text.

   /   /2019

                                   Miss. ZE Mbhele

As the candidate's supervisor I agree to the submission of this dissertation.

   /   /2019

                                   Prof. THT Coetzer

Declaration-Plagiarism

I, Ziphezinhle Elaine Mbhele, declare that:

1. The research reported in this dissertation, except where otherwise indicated, is my original research.
2. This dissertation has not been submitted for any degree or examination at any other university.
3. This dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   a. Their words have been re-written but the general information attributed to them has been referenced
   b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.
5. This dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the references sections.

     /   /2019

                                         _____
                                         Miss. ZE Mbhele

Abstract

African trypanosomiasis is a disease caused by protozoan parasites i.e. *Trypanosoma* spp. in livestock and humans and affects 37 sub-Saharan countries. Animal African trypanosomiasis (AAT) is known as nagana and African human trypanosomiasis (HAT) as sleeping sickness. *Trypanosoma congolense*, *T. vivax* and *T. brucei brucei* cause AAT which is an economic burden and hampers agricultural development in Africa. The parasite escapes the host's immune response by switching the genes coding for the variable surface glycoproteins, resulting in new variable antigen types. This has made it unlikely to develop a vaccine against the disease and therefore many studies now focus on non-variant trypanosome antigens as potential diagnostic and drug targets.

Trypanosomal cysteine proteases, such as the cathepsins B and L, have been identified and validated as potential diagnostic targets. They are expressed throughout the parasite life cycle and are essential for the survival of the parasite. The mammalian host produces an antibody response against trypanosomal cysteine proteases which do not affect the survival of the parasite, however, the antibodies are believed to play a role in trypanotolerance by neutralising the effects of the enzyme.

Antigen-based ELISA is a good tool for accurate diagnosis of AAT, but also relies on good antibodies. The overall aim of this study was to produce antibodies against the cathepsin-L-like protease of *T. b. brucei* (*Tbb*CATL) which can be used for specific diagnosis of *T. b. brucei* infections. This included polyclonal antibodies as well as single chain variable fragments (scFvs) using phage display. The protease, *Tbb*CATL, was recombinantly expressed in *E. coli* for the first time as a 61 kDa protease (including the GST tag) using the pGEX-4T expression vector. The homologues from *T. congolense* (*Tco*CATL) (29 kDa) and *T. vivax* (*Tvi*CATL) (28 kDa and 32 kDa) were also recombinantly expressed using the *P. pastoris* yeast expression system and were shown to hydrolyse the Z-Phe-Arg-AMC substrate and to be inhibited by E-64. Antibodies against whole *Tbb*CATL were produced in chickens and together with anti-*Tco*CATL peptide antibodies and anti-*Tvi*CATL antibodies, were evaluated in a western blot to determine possible cross-reactivity. Whereas the anti-*Tbb*CATL antibodies were specific for *Tbb*CATL, the anti-*Tco*CATL peptide and anti-*Tvi*CATL antibodies cross-reacted with *Tbb*CATL. The production of scFvs was optimised using *Tvi*CATL as the panning antigen against the Nkuku® phage display library. The *Tvi*CATL-specific phages were enriched after the fourth round of panning and a total of seven clones gave high signals when analysed by monospecific ELISA. Future work will include recombinant expression and purification of the selected *Tvi*CATL-specific scFvs for testing in diagnostic assays as well as panning with the *Tbb*CATL antigen. This study laid the groundwork for evaluating *Tbb*CATL as a diagnostic target for AAT

Acknowledgments

"To God our saviour, who alone is wise, be all glory, dominion and power, both now and forever. Amen"-Jude 1:25. I would like to thank the Lord almighty who has and continues to carry me through all things.

A big thank you to Prof. Theresa. H.T Coetzer for giving me this wonderful opportunity of doing my MSc under your supervision and guiding me through every step. It was such an honour to work with you.

To Dr Lauren Eyssen, thank you for your mentorship and advice throughout this project and ensuring organisation in our lab. Dr Rob Krause and Dr Murtala Isah, thank you for all the help and support you gave me since honours.

To all my colleagues, who became my family: Lucky, Faiaz, Ephraim, Amukelani, Nomusa, Chanelle, Thando, Nxalati, Pretty and Sindy. Thank your all your advice, support and encouragement and for the fun times in and out of the lab. Without you I wouldn't have survived.

Thank you to my fellow post graduated and friends: Ntethelelelo, Nomfundo, Andile, Lisa, Abdul, Sno, Alex, Mark, Sthabiso, Sheldon, Ntobe and Mbali.

To the entire staff in the school of Life science, especially the biochemistry department for ensuring a smooth functioning, safe and clean environment for us to work in.

A special thank you to my siblings, Nompumelelo, Snethemba, Sanelisiwe and Ndumiso for all the love and making my journey through life bearing. I love you so much.

I dedicate this dissertation to my mother, the love of my life.

Table of Contents

List of Figures

List of Tables

Abbreviations

2xYT	2 x yeast tryptone
AAT	animal African trypanosomiasis
AMC	7-amino-4-methylcoumarin
AT	African trypanosomiasis
BiP	immunoglobulin binding protein
Bis	N,N'-methylenebisacrylamide
bp	base pair
BSA	bovine serum albumin
BSF	bloodstream form
CATB	cathepsin B
CATT	card agglutination test
CATL	cathepsin L
CDR	complementary determining regions
CFT	complement fixation test
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dsDNA	double stranded deoxyribonucleic acid
DTT	dithiothreitol
E-64	L-*trans*-epoxysuccinyl-leucylamido(4-guanidino) butane
ELISA	enzyme-linked immunosorbent assay
Fab	fragment antigen binding
FAO	Food Agriculture Organisation
Fc	fragment crystallisable
Fv	fragment variable
g	relative centrifugal force
GST	glutathione S transferase
HAT	human African trypanosomiasis
HRPO	horseradish peroxidase
IFAT	indirect immunofluorescence antibody
Ig	immunoglobulin
IgG	immunoglobulin G
IgM	immunoglobulin M
IgY	immunoglobulin Y
IPTG	isopropyl β-D-1 thiogalactopyranoside
ISG	invariant surface protein
kDa	kilo-Dalton

kDNA	kinetoplast deoxyribonucleic acid
LAMP	loop-mediated isothermal amplification
MEC	molecular exclusion chromatography
min	minute/s
Mr	relative molecular mass
mRNA	messenger ribonucleic acid
NECT	nifurtimox- eflornithine combination therapy
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCF	procyclic form
PCR	polymerase chain reaction
PEG	polyethylene glycol
RNA	ribonucleic acid
rpm	rotations per minute
RT	room temperature (22 ± 2°C)
scFv	single chain variable fragment
SDS	sodium dodecyl sulfate
ssDNA	single stranded deoxyribonucleic acid
TAE	Tris-acetate-EDTA
TBS	tris-buffered saline
TEMED	N,N,N',N'-tetramethyl ethylenediamine
TPP	three phase partitioning
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
tRNA	transporter ribonucleic acid
Trx	thioredoxin
TYE	tryptone yeast extract
VSG	variant surface glycoprotein
VAT	variant antigen type
WHO	World Health Organisation
X-gal	5-bomo-4-chloro-3-indyle-β-D-galactopyranoside

Chapter 1

Literature Review

1.1 Introduction

African trypanosomiasis (AT) is a disease caused by flagellated protozoan parasites i.e. *Trypanosoma* spp. in livestock and humans and affects 37 sub-Saharan countries. African animal trypanosomiasis (AAT) is known as nagana and human African trypanosomiasis (HAT) as sleeping sickness. There are approximately 30 species of tsetse flies (*Glossina* spp.) inhabiting the humid regions of Africa which are responsible for transmitting both HAT and ATT causing trypanosomes (Muhanguzi *et al.*, 2017). The prevalence of the trypanosome parasite correlates with the distribution of the tsetse fly vector which covers 10 million km^2 land in sub-Saharan Africa (Fig. 1.1) (Yaro *et al.*, 2016). The tsetse fly is found between 14°N and 29°S of the equator where the environment is suitable for their survival and reproduction (Mitchell, 2018).

Sleeping sickness is caused by two *Trypanosoma brucei* sub-species namely *Trypanosoma brucei gambiense* which causes the chronic form in west and central Africa, and *T. b. rhodesiense* which causes the acute form in east and southern Africa (Brun *et al.*, 2010). This disease occurs in two clinical stages. The first stage is known as the haemo-lymphatic stage where the parasites invade the lymphatic system and bloodstream. After some time, the parasites cross the blood-brain barrier and invade the central nervous system, this is the second clinical stage known as the meningo-encephalytic stage (Brun *et al.*, 2010). Nagana, on the other hand, is caused by *T. congolense*, *T. vivax* and to a lesser extent *T. b. brucei.* These species invade the bloodstream causing anaemia (Noyes *et al.*, 2009). Whereas other species develop in the vector midgut, *T. vivax* develops in the mouthparts of the insect vector where it undergoes a complete life cycle. Due to this, *T. vivax* can be mechanically transmitted by other haematophagous flies such as horseflies (*Tabanus* spp.) and stable flies (*Stomoxys* spp.) and is consequently also found in parts of Asia and South America (Giordani *et al.*, 2016; Radwanska *et al.*, 2018).

Another important trypanosoma species is *T. cruzi* which causes American trypanosomiasis, known as Chagas diseases in humans. There are more than 100 blood sucking *Triatominae* species that can serve as vectors for this parasite. While transmission of *T. b. gambiense* and *T. b. rhodesiense* occurs via the saliva of vectors during a blood meal, *T. cruzi* is deposited with parasite-laden faeces of the vector during a blood meal. Infected individuals develop cardiac and gastrointestinal problems associated with chronic symptoms of Chagas disease (Kowalska *et al.*, 2011).

The Food and Agriculture Organisation (FAO) of the United Nations declared trypanosomiasis a major threat to the economy and agricultural production on the African continent (Mattioli *et al.*, 2004). At least three million cattle die annually, affecting the livelihood of local farmers who depend on cattle as a source of income, meat, milk and manure (Giordani *et al.*, 2016). To date, there is no conventional vaccine for trypanosomiasis due to the continuously changing surface coat of the parasites, hence this disease is controlled mainly by trypanocidal drugs. The effect of these trypanocides, however, is limited by several factors, the major factor being the emergence of drug-resistant strains. Due to this, many research studies have been aimed at understanding drug-susceptible strains for appropriate drug prescription and administration, as well as host-parasite interactions to provide the understanding of the disease and drug resistance mechanisms (Sutcliffe *et al.*, 2014). Furthermore, the FAO has implemented quality control checks to improve the quality of veterinary drugs (Giordani *et al.*, 2016).

Figure 1.1: Distribution map of tsetse flies and cattle in Africa. Accessed from http://www.irinnews.org/news/2009/05/12/tsetse-fly-costs-agriculture-billions-every-year.

1.2 Classification of trypanosomes

Trypanosomes belong to the genus *Trypanosoma*, family Trypanosomatidae, order Kinetoplastida (due to the presence of a kinetoplast) and class Sacomastigophora (Hoare, 1972). The genus is further divided into two groups namely stercoraria and salivaria (Fig. 1.2). In the stercoraria group, the parasite develops in the intestine of the invertebrate vector and transmission to the vertebrate occurs via faeces (e.g. *T. cruzi* which is responsible for Chagas disease). The parasites of the salivaria group develop in the gut of the vertebrate host and move towards the salivary gland where they develop into the form that is infectious to the vertebrate host. Included in this group are subgenera *Nannomonas, Duttonella*, *Trypanozoon* and *Pyconomonas*. Subspecies *T. congolense* and *T. vivax* belong to subgenera

Nannomonas and *Duttanella* respectively and are pathogenic to livestock (Table1.1). Subgenera *Trypanozoon* includes *T. brucei* with three sub-species; *T. b. brucei* (infects livestock but is not pathogenic), *T. b. rhodesiense* and *T. b. gambiense* which cause sleeping sickness in humans (Majiva *et al.*, 1986). The *Trypanozoon* species *T. equiperdum* and *T. evansi,* are transmitted by biting insects and infect camels and horses in northern Africa and parts of Asia (Table 1.1) (Radwanska *et al.*, 2018).

1.3 Morphology and life cycle of trypanosomes

The trypanosome is a unicellular organism which varies from 8 to 50 μm in size. This parasite is unique among the flagellated protozoans in that the flagellum is attached to the cell body (Fig. 1.3) producing waves from the ends of the body, resulting in a wriggling movement observed in culture media or blood smears (Schuster *et al.*, 2017). The parasite is characterised by a kinetoplast, which is a compact disc of mitochondrial DNA known as kinetoplast DNA (kDNA). The kDNA contains two types of topologically interlocked DNA circles i.e. maxi circles and mini circles. Maxi circles (22 kb) are found in dozens per cell and contain conserved sequences that code for RNA and a few mitochondrial proteins. To form functional mRNA, the maxi circles require high fidelity editing which is controlled by a guide mRNA encoded in the high copy number mini circles (1 kb) which have variable sequences (Liu *et al.*, 2005).

The position of the kinetoplast changes during the different life cycle stages. During the trypomastigote stage (in the host) it is located at the posterior end of the organism and during the epimastigotes stage (in the insect vector) it is located at the anterior end. The mitochondrion extends towards the anterior end and does not contain cristae during the bloodstream form therefore, mitochondrial respiration does not occur. In this stage, energy generation occurs in the glycosomes via glycolytic reactions. On the other hand, the mitochondrion is highly active during the procyclic form as energy generation is dependent on the electron transport chain in the absence of blood glucose (Matthews, 2005).

The nucleus contains the DNA responsible for the synthesis of enzymes and proteins of the cell. Beneath the cell membrane are tightly packed microtubules which limit access into the cell, thus the flagellar pocket (specialised invagination) is the only site for endocytosis and exocytosis (Parsons, 2004; Matthews, 2005). Material internalised for nutritional or immune invasion purposes enter the parasite by endocytosis and are transported via endosome compartments to the lysosomes where they are proteolytically degraded (O'Brien *et al.*, 2008).

Phylum

Protozoa

Order

Kinetoplastida

Family

Trypanosomatidae

Genus

Crithidia   *Trypanosoma*   *Leishmania*

Group

Salivaria   Stercoraria

Subgenus

Trypanozoon   *Nannomonas*   *Pycnomonas*   *Duttonella*   *Herpetosoma*   *Megatrypanum*   *Schizotrypanum*

Species

Subspecies

Trypanosoma evansi

T. equiperdum

T. brucei:

  T. b. brucei

  T. b. gambiensis

  T. b. rhodesiense

Trypanosoma godfreyi

T. simiae

T. congolense

Trypanosoma suis

Trypanosoma vivax

  T. v. viennei

  T. uniforme

Trypanosoma lewisi

T.musculi

Trypanosoma theileri

Trypanosoma cruzi

  T.c.marinkellei

  T. rangeli

Figure 1. 2 : Classification of trypanosomes (Stevens and Gibson, 1999; Stevens and Brisse, 2004; Baral, 2010). This review focuses on the underlined species.

Table 1.1: Animal infective *trypanosoma* species and geographic areas where they occur (Giordani *et al.*, 2016; Cooper *et al.*, 2017)

Trypanosoma spp.	Affected animals	Geographic distribution
T. congolense	Cattle, sheep, goats, pigs, horses, dogs	Tsetse regions of Africa
T. vivax	Cattle, sheep, goats, horses, camels	Africa, central and south America, west Indies
T. brucei brucei	Cattle, dogs, camels, horses	Tsetse regions of Africa
T. simiae	Sheep, goats, camels, horses, pigs	Tsetse regions of Africa
T. equiperdum	Horses, donkeys, mules	Africa, Asia, middle east
T. evansi	Cattle, horses, water buffalo, dogs	Asia, Africa, South and Central America
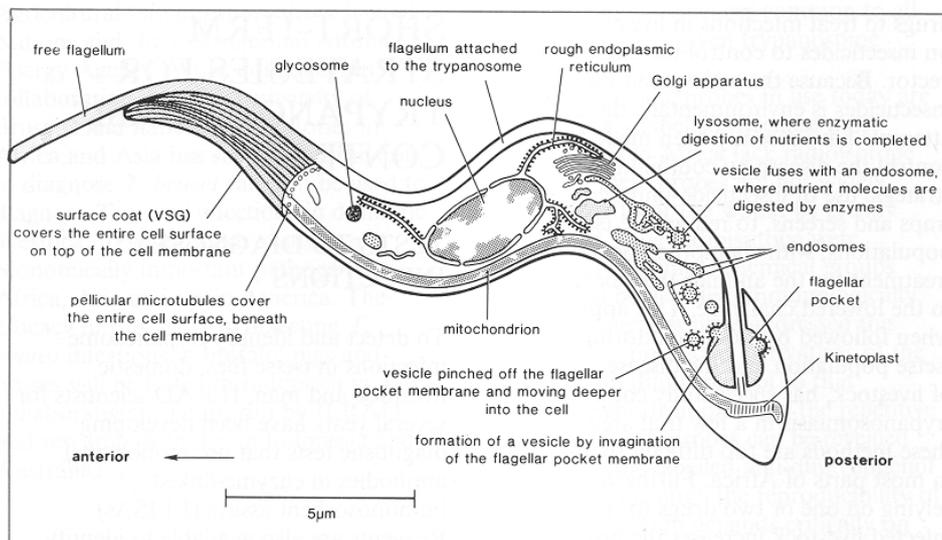
Figure 1.3: Diagrammatic representation of the morphology of a trypanosome, *T. brucei* showing its major organelles. Accessed from http://agsavet.blogspot.com/2011/12/trypanosomiasis.html

Trypanosomes initially multiply within the midgut of the insect vector before migrating to the mouthparts as the mammalian infective metacyclic form (Fig. 1.4; stages 1 to 5). The metacyclic parasite develops in the proboscis in the case of *Nannomonas* and the salivary

glands in the case of *Trypanozoon.* During a blood meal, the insect vector injects the metacyclic parasites into the skin of the mammalian host from where they enter the blood via the lymphatic system. In the early stages, the parasites are long, slender and divide rapidly by binary fusion causing an increase in parasitaemia. During peaks of parasitaemia, the stumpy form is established, does not divide and is pre-adapted for survival in the vector after it is taken up in a blood meal. In the vector midgut, the stumpy form differentiates into the procyclic form. The parasite moves through the midgut to the proventriculus and differentiate into epimastigote and finally the metacyclic form in the salivary glands (Fig. 1.4) (Sunter and Gull, 2016).
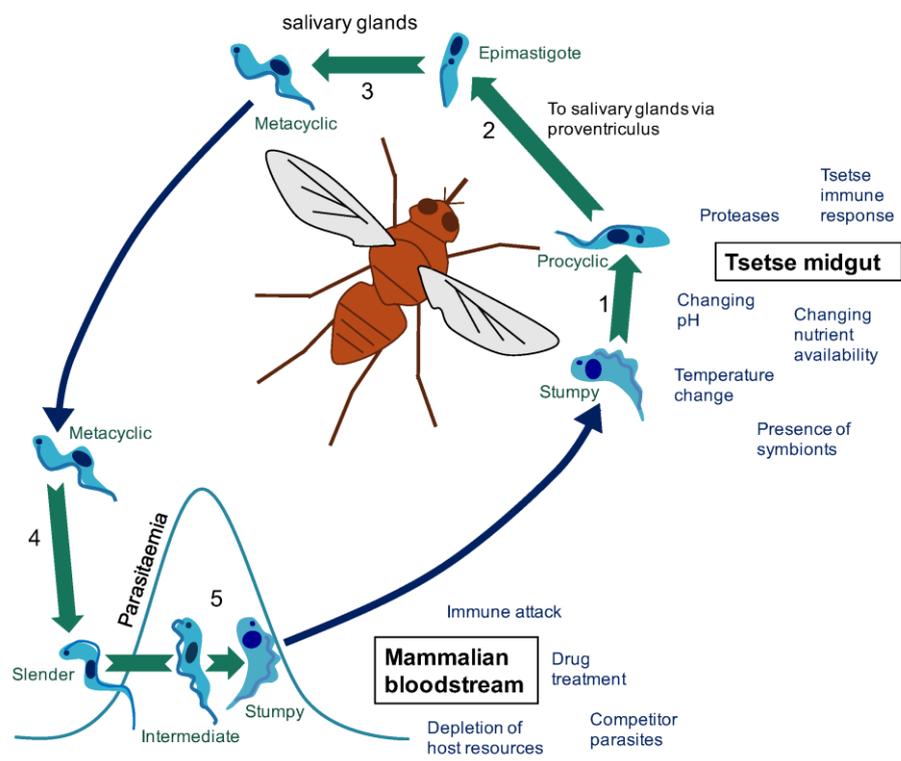
Figure 1.4: Schematic representation of the life cycle of *T. brucei* in the tsetse and mammalian hosts. (Silvester *et al.*, 2017).

1.4 Antigenic variation of trypanosomes

A feature unique to trypanosomes is the replaceable cover coat made up of variable surface glycoprotein (VSG) which specifies the variable antigen type (VAT) (Ponte-Sucre, 2016). There are more than 10^7 VSG genes encoded in the parasite's genomic repertoire. Transcription of the VSG mRNA occurs from one of many telomeric bloodstream form sites while the other sites are silenced, therefore only one gene at a time is expressed on the surface of the parasite at any time (Glover *et al.*, 2013). During mammalian infection, the VSG

generates a host immune response which eliminates the majority of the parasites, and only a few slender cells survive by switching the expressed VSG, resulting in a new wave of parasitaemia. The stumpy cells on the other hand, do not switch their VSG but enter cell cycle arrest thus cannot generate new variants. These cells however, can rapidly eliminate anti-VSG antibodies by endocytosis, causing a delay in cell lysis. Due to this, their survival is prolonged which increases the chances of transmission (Silvester *et al.*, 2017). The VSG covers about 95% of the parasite's surface thereby shielding other surface proteins from the host's immune response (Schwede *et al.*, 2015). Furthermore, VSGs can remove specific host anti-VSG antibodies from the cell surface thereby reducing the effect of the host's antibody response (Engstler *et al.*, 2007).

1.5 Clinical symptoms of trypanosomiasis

In humans, the early stages of trypanosome infection are known as the haemolymphatic phase during which the parasites are restricted to the blood and lymph system (Stijlemans *et al.*, 2015). Following an insect bite, infected patients may develop a local skin reaction caused by an inflammatory response. Early symptoms include fever due to successive waves of parasitaemia, headaches, weight loss, itching, weakness and pain in the joints. Immune responses also result in enlarged lymph nodes and spleen. The disease then develops to the meningo-encephalitic stage where the parasite crosses the blood-brain barrier and enters the central nervous system and cerebrospinal fluid, resulting in confusion, disturbed sleeping pattern, lethargy and a coma. The terminal phase of the disease is characterised by seizures and heart failure (Sternberg, 2004). If untreated, infected patients may die within months if infected with *T. b. rhodesiense* or within years if infected with *T. b. gambiense* (Steverding, 2008).

Cases of animal trypanosomiasis are usually chronic but acute cases also occur leading to death within a week of infection. Similar to HAT, there is usually swelling at the site of the insect bite which may remain unnoticed. Infected animals develop primary symptoms such as fever, weight loss, lymphadenopathy and anaemia (Steverding, 2008). Anaemia is diagnosed by the low packed volume of red blood cells. These cells undergo lysis due to the innate immune response or by haemolysins released by the parasite (Cooper *et al.*, 2017). Further symptoms may include decreased milk production, neurological signs, cardiac lesions, diarrhoea, keratitis, loss of appetite, emaciation, weight loss and paralysis (Steverding, 2008).

1.6 Diagnosis of trypanosomiasis

An understanding of the epidemiology of trypanosomiasis and management of the disease depends on accurate diagnosis and distinction between the causative species. Clinical signs

and symptoms are the initial diagnosis of trypanosomiasis. In regions where HAT is endemic, presentation of clinical signs in patients suggests possible infection (Kennedy, 2013). In the case of AAT endemic areas, anaemia in cattle is usually an indication of infection (Auty *et al.*, 2015). Trypanosomes may also cause chancre which can be detected several days after a tsetse bite, however, not all trypanosome strains cause chancre. Additionally, the symptoms may be mistaken with that of other diseases such as diabetes or theileriosis in cattle thus making clinical diagnosis less reliable (Luckins and Gray, 1978). In addition to clinical signs, serological, parasitological and molecular tests are usually required for trypanosomiasis diagnosis.

1.6.1 Parasitological diagnosis for trypanosomiasis

Parasitological tests are based on standard microscopic detection of trypanosomes in blood samples or lymph node smears. These were the first parasitological tests done in endemic areas as parasites can be easily detected under a light microscope based on their morphology and light microscopy is still widely used for diagnosis of AAT (Dabo and Maigari, 2017). Thick blood smears can be used to detect the parasite but does not distinguish between trypanosome species. Thin smears, on the other hand, can be used to distinguish between the species based on morphology and wet mounts distinguish based on motility. *T. vivax* is characterised by rapid movements across the microscope slide while *T. brucei* moves in one place. *T. congolense* attaches onto red blood cells and moves in one place (Dabo and Maigari, 2017).

The sensitivity of the technique was improved by the removal of the red blood cells followed by concentration of the parasites by centrifugation prior to microscopic analysis. Later, a more sensitive technique was developed, the mini anion-exchange centrifugation technique (mAECT) based on the removal of blood cells from the test blood sample through anion exchange chromatography. This technique was reported to detect approximately 100 trypanosomes per mL of blood and worked best for *T. brucei* and *T. evansi* spp. (Büscher, 2002; Büscher *et al.*, 2009; Büscher *et al.*, 2017).

The stage of HAT infection is determined by microscopic examination of the cerebrospinal fluid (CSF) obtained through a lumbar puncture. The presence of the parasite and an increase in white blood cell count are an indication of stage two HAT (Dabo and Maigari, 2017). Staging of HAT is crucial for treatment purposes as the second stage requires medication that can cross the blood-brain barrier. The sensitivity of this technique allows the identification of non-HAT patients to ensure that they are not unnecessarily treated for second-stage HAT (Chappuis *et al.*, 2005). However, false negatives can occur with this method of diagnosis,

particularly in *T. b. gambiense* infections where the parasite numbers are sometimes below the detection limit.

1.6.2 Molecular diagnosis for trypanosomiasis

Molecular diagnostic tests are based on the detection of the parasite's DNA. Understanding the trypanosomes at a genomic level has led to the development of specific primers which can be used in a polymerase chain reaction (PCR) and loop-mediated isothermal amplification (LAMP) (Solano *et al.*, 1999). The latter has advantages over PCR in that: partially processed or non-processed DNA template can be used therefore DNA extraction may not be required. It also uses a simpler heating device that maintains temperature and has a shorter amplification time (Njiru *et al.*, 2011). Molecular diagnostic tests have been improved by the identification of DNA based markers that differentiate between trypanosome species and subspecies (Deborggraeve and Büscher, 2010).

In the diagnosis of HAT, the molecular differentiation between *T. b. gambiense* and *T. b. rhodesiense* is based on two gene markers, i.e. *T. b. gambiense* specific glycoprotein which is only found in *T. b. gambiense* and serum-resistance-associated protein specific for *T. b. rhodesiense* (Gibson *et al.*, 2002). The PCR system has also been used for the diagnosis of Chagas disease and these are based on the satellite DNA of *T. cruzi* (Abras *et al.*, 2018).

Both PCR and LAMP technology have also been used to identify the three AAT causative species. Specific primers for *T. congolense*, based on ribosomal P0 subunit protein, and *T. vivax*, based on satellite DNA, have been designed and used to diagnose AAT (Kuboki *et al.*, 2003; Njiru *et al.*, 2011).

It is evident that diagnosis depends on the identification of specific antigens. Both the DNA based techniques are useful in that they can directly detect the pathogen, but the parasitaemia can vary. Additionally, most of these molecular tests are not applicable to the resource-poor field settings therefore, appropriate serological tests are needed (Auty *et al.*, 2015).

1.6.3 Serological diagnostic methods for trypanosomiasis

Trypanosome infection can also be indicated by the presence of specific antibodies in the blood, plasma or serum of the infected hosts. The serological tests are often based on VSGs because these proteins are highly immunogenic and expressed in higher levels compared to other trypanosome proteins. The sensitivity of antibody detection tests depends on the type of antigen(s) used. Serological tests detect antibodies three to four weeks after infection (Mugnier *et al.*, 2016).

1.6.3.1 Serological diagnosis of human African trypanosomiasis

Currently, there is a serodiagnostic test for *T. b. gambiense* infections available, the card agglutination test for trypanosomiasis (CATT), which detects specific antibodies against VSG variant LiTat 1.3 in the blood serum of infected individuals. This test has also been used to detect *T. b. evansi* infections by detection of RoTat 1.2 VSG (Ngaira *et al.*, 2005). The card agglutination test, however, is associated with a few limitations which include a short shelf life and a requirement for electric power, therefore, cannot be used in rural areas (Jamonneau *et al.*, 2015).

Rapid diagnostic tests for *T. b. gambiense* infection include HAT- Sero-K- SeT and SD Bioline HAT 1.0 which detect antibodies against variable antigen types LiTat 1.3 and 1.5 VSGs (Büscher *et al.*, 2013). The advantage of these diagnostic tests is that they comply with the ASSURED (affordable, sensitive, specific, user-friendly, rapid and robust, equipment-free and deliverable to end users) criteria. Unfortunately, the *T. b. rhodesiense* strain does not express the LiTat 1.3 VSG, consequently, diagnosis still relies on clinical features and blood smears are used to confirm the diagnosis. Efforts towards the development of rapid diagnostic tests for both *T. b. gambiense* and *T. b. rhodesiense* are still on-going (Büscher *et al.*, 2017). Although CATT and rapid diagnostic tests are useful in the screening and diagnosis of *T. b. gambiense* infections, these tests are not 100% accurate and give false negatives when the disease prevalence is low (0.1%) (Jamonneau *et al.*, 2015).

Other antibody detection tests include the complement fixation test (CFT), indirect fluorescent antibody (IFAT) test and enzyme-linked immunosorbent assay (ELISA) for both HAT and AAT diagnosis. ELISA has been used in epidemiological surveys to detect anti-trypanosome antibodies and is considered a reliable and easy immunological diagnostic method. It is sensitive, specific and can be applied to dry blood spots. However, the technique is based on either whole parasite or crude lysate as the antigen, making it difficult to standardise (Mugnier *et al.*, 2016).

1.6.3.2 Serological diagnosis of animal African trypanosomiasis

There are non-variant molecules which have been identified as diagnostic targets and are immunogenic in mammalian hosts. A mitochondrial heat shock protein (HSP70) from *T. congolense,* a mammalian immunoglobin binding protein (BiP) homologue has been proposed as a potential diagnostic target. Immunoglobin binding protein homologues are also found in *T. vivax* and *T. brucei* (Boulangé et al., 2002). An inhibition ELISA based on the HSP70/BiP was able to detect primary infections in cattle challenged with *T. congolense* and *T. vivax* while further studies are required to validate this antigen as a diagnostic target for *T. brucei* (Bossard

et al., 2010). *T. vivax* infections have also been successfully detected in cattle using a lateral flow test based on invariant surface glycoprotein (ISG). Recombinant ISG was able to distinguish infected sera from non-infected sera in an immunoreaction, with equal sensitivity to an ELISA test (Fleming *et al.*, 2016). A tandem repeat protein GM6 and cathepsin B-like cysteine protease were also identified as potential serodiagnostic targets for *T. vivax* and *T. congolense* infections respectively. However, the cathepsin B-like cysteine protease shares epitopes with the *T. vivax* homologues which results in cross-reactivity in a rapid diagnostic test (Boulangé *et al.*, 2017). *T. vivax* recombinant cathepsin L-like cysteine proteases (*Tvi*CATL) was evaluated as a diagnostic antigen and could discriminate between *T. vivax* infected and non-infected cattle sera (Eyssen *et al.*, 2018).

Because anti-trypanosomal antibodies persist in the blood, ELISA does not specify whether the infection is current or whether the animals were cured of previous infection (Van den Bossche *et al.*, 2000). Therefore, an alternative to antibody-based detection ELISA is antigen-based detection ELISA. The latter detects specific antigens in the blood of infected hosts which is an indication of current infection. This approach was first used to diagnose Chagas disease using polyclonal antibodies. Polyclonal antibodies, however, have low specificity, thus the use of species-specific monoclonal antibodies was developed. These were successful in the detection of *T. b. rhodesiense* and *T. b. gambiense* infections in humans. Monoclonal antibodies have also been used in antigen-capture sandwich ELISA to detect *T. congolense*, *T. vivax* and *T. brucei* in the blood of infected cattle and goats. However, this type of test was also shown to give false negatives in acute or early infections. This is because the antibodies are directed at internal antigens which are only released after lysis of the parasite. As a result, antigen ELISAs are combined with parasitological tests for an accurate diagnosis. Furthermore, the use of recombinant antigens is a promising tool for the detection of trypanosome antibodies (Mugnier *et al.*, 2016).

1.7 Control measures for trypanosomes

Accurate diagnosis followed by appropriate treatment are important for effective control of AT. This is a major challenge in the case of HAT as it is often under-diagnosed due to limited accuracy in diagnostic methods, insufficient numbers of skilled staff and the presence of an animal reservoir host, particularly for *T. b. rhodesiense* HAT. Control by surveillance has been identified as a useful technique. This is where a mass screening is done in areas identified as high risk, followed by treatment upon case detection (Büscher *et al.*, 2017). Though the reported reduction in HAT cases in the past decade, there is limited access to remote areas and people are often reluctant to participate in the repeated screening activities which are a major challenge with this technique, making it less effective (Franco *et al.*, 2014).

Tsetse vector control is considered a desirable method of AT control. The Pan African tsetse and trypanosome eradication campaign (PATTEC) together with the African Bank Development have partnered in an attempt to eliminate the tsetse vector by making funding available for this initiative (Holt *et al.*, 2016). Since the 1980s vector control methods have included selected bush clearing, sequential aerial spraying, mobile and stationary traps treated with insecticides, sterile insect technique as well as treatment of cattle with insecticides. These are done by using either a field-by-field approach where the goal is to temporarily alleviate AT burden in a localised area or an area-wide approach where the aim is to eliminate the entire tsetse population in a geographical area (Vreysen *et al.*, 2013). Although there are government charities and funding towards this initiative, a number of communities who are affected by the disease do not have enough resources for vector control and these programmes do not always reach them (Holt *et al.*, 2016).

1.7.1 Treatment of trypanosomiasis

Vector control methods are expensive, especially when applied in large scale, therefore, most African farmers opt for trypanocidal drugs in the case of AAT (McDermott and Coleman, 2001; Giordani *et al.*, 2016). Trypanocides are divided into two categories i.e. curative trypanocides which kill the parasite in the blood of infected animals but are not always 100% effective, and chemoprophylactic drugs which kill the parasites and remain in the blood to prevent new infections (Desquesnes *et al.*, 2013).

The use of trypanocides in the chronic phase of AAT does clear parasitaemia but the symptoms may take longer to subside. The commonly used curative drug is diminazene aceturate and is used to treat infections in cattle, sheep and goats. Diminazene is administered in high doses through intramuscular injections. This drug is active against *T. evansi*, *T. congolense T. brucei* and *T. vivax* and has few side effects (Desquesnes *et al.*, 2013). Other drugs include isometamidine chloride, melarsomine dihydrochloride, suramin and quinapyramine which are used as prophylactic drugs. These drugs are often administered by the farmers themselves with no supervision which has contributed to the emergence of drug resistance, a major challenge since there is a limited number of drugs commercially available (Giordani *et al.*, 2016).

There are five drugs available for the treatment of HAT. The choice of drugs depends on the causative species and stage of the disease. Pentamidine and suramin are used to treat the first stage of HAT. Pentamidine is more effective against *T. b. gambiense* HAT while efficacy is limited in *rhodesiense* HAT (Simarro *et al.*, 2012). Pentamidine is administered intramuscularly or intravenously. Suramin, on the other hand, is effective against both HAT

causing species but is only used to treat *T. b. rhodesiense* HAT due to the high risk of allergic reactions caused by the rapid killing of microfilaria (Kennedy, 2013). Early infections are easier to treat than second stage diseases because the latter requires drugs that cross the blood-brain barrier, which may be toxic to the patient (Büscher *et al.*, 2017).

Melarsoprol, nifurtimox and eflornithine are used to treat the second stage of HAT. Melarsoprol is used to treat *T. b. rhodesiense* infections over a 10-day period of intravenous injections. Eflornithine is effective against *T. b. gambiense* infections and is administered intravenously over 14 days. Nifurtimox is used to treat Chaga's and is not licenced in Africa but is used as an off-label treatment in west Africa, which requires authorisation by national authorities (Büscher *et al.*, 2017). These drugs are associated with several issues of toxicity and resistance, therefore, efforts are continually being made for better and safer use (Babokhov *et al.*, 2013). Additionally, there is limited availability of the drugs and they are often administered in large doses. A combinatory administration of the drugs may be a solution to these drawbacks. An example of this is nifurtimox-eflornithine combination therapy (NECT) which was found to have a higher curative rate and lower death rates compared to melarsoprol and eflornithine monotherapies (Büscher *et al.*, 2017). A new oral drug, fexinidazole was shown to be effective against second stage *T. b. gambiense* HAT. This drug has a higher efficacy compared to NECT and is available at a low cost. However, it is also associated with adverse side effects such as vomiting and psychiatric disorders (Chappuis, 2018).

The development of trypanocidal drug resistance is becoming a major problem in the treatment of trypanosomiasis. This is characterised by the loss of sensitivity to trypanocides even at doses recommended and under proper veterinary practises. These drugs have been in use for over 60 years and resistance has been reported in many areas. Resistance to trypanocidal drugs could be innate, which occurs naturally in hosts that have not been previously exposed to a drug, acquired resistance, which is a result of previous exposure or due to mutation in the parasite's genome. Understanding the mechanism by which trypanosomes take up drugs is important for the development of new chemotherapeutic strategies (Melaku and Birasa, 2013).

Drug resistance has been reported for one or more of the AAT drugs in more than 17 sub-Saharan countries. Multiple resistance has also been reported for diminazene and isometamidine chloride used for AAT (Giordani *et al.*, 2016). In the case of HAT treatment, the introduction of NECT reduced the chances of drug resistance (Büscher *et al.*, 2017).

1.7.2 Trypanotolerance

Trypanotolerance in cattle is defined as the ability to survive and remain productive during trypanosomal infections. It was first observed when cattle of taurine origin, N'Dama cattle *(Bos taurus)*, were found to be less susceptible to trypanosomiasis (Murray *et al.*, 1981). This field observation was confirmed in a comparative study between N'Dama and Boran cattle challenged with *T. congolense* which showed that N'Dama cattle remained productive, gained weight and had better control over parasitaemia and anaemia than the Boran breed (Paling *et al.*, 1991).

Since these findings, many studies have been directed towards understanding and identifying the genes associated with trypanotolerance and discovering processes that could help with the conference of these genes. A study by Hanotte *et al.* (2002) revealed that F2 calves derived from Boran and N'Dama contained five genes from Boran grandparents which were involved in trypanotolerance. This suggested that crossbreeding may be a useful approach not only to improve the trypanotolerance of N'Dama cattle but also in the control of trypanosomiasis (Naessens, 2006).

Additionally, in the early stages of trypanosome infection, N'Dama cattle were found to produce high levels of IgG reactive with *T. congolense* and *T. b. brucei* lysosomal cysteine protease which was not the case in susceptible cattle. This cysteine protease is associated with the pathogenicity of the trypanosome parasite and has also been reported in other parasites (discussed in Section 1.8). Despite the higher levels of antibody response to other proteins in trypanotolerant cattle compared to susceptible cattle, antibody response to the cathepsin L-like cysteine protease was regarded a marker of trypanotolerance (Orenge *et al.*, 2012; Berthier *et al.*, 2015).

1.8 Proteases

Proteases are enzymes which hydrolyse peptide bonds of proteins. Although all proteases cleave proteins, they differ in many characteristics including size, localisation, quaternary structure and the mechanism used to hydrolyse the substrate. These enzymes can be further classified according to the active site groups involved in the catalytic mechanism into seven classes i.e. serine, cysteine, aspartic, threonine, glutamic and metalloproteases as well as asparagine peptide lyases (Otto and Schirmeister, 1997). Among the proteases, cathepsins play a major role in proteolysis, particularly the aspartic, serine, and cysteine cathepsins (Lecaille *et al.*, 2002). The majority of the cathepsins are cysteine proteases such as cathepsin B, C, F, H, K, L, S, V, X and W. Cathepsins D and E are aspartic cathepsins, and cathepsins A (also known as carboxypeptidase Y) and G are serine cathepsins (Turk and Turk, 2009).

1.8.1 Cysteine proteases

The first cysteine (thiol) protease was isolated and characterised from a *Carica papaya* fruit, thus named papain. Since the discovery of papain, cysteine proteases that have a DNA sequence like papain are called "papain-like". Cysteine proteases are found in all living organisms, have a molecular mass of ~21-30 kDa and show optimum activity at pH 4 to 6.5 (Turk *et al.*, 1998). Sequences for several cysteine proteases have been identified in the human genome and parasitic organisms. These include the main groups of cysteine proteases: clan CA, clan CD and caspases. Clan CA is further subdivided into the C1 family (cathepsins B and L) known as papain-like cysteine proteases due to their similarities in structure, and C2 family (calpain). Most studies have been focused on papain-like cysteine proteases cathepsins B and L because they are found in a wide range of parasitic organisms. The mechanism of these proteases is well understood, which is an advantage in the process of drug discovery. Furthermore, it is believed that due to their similarities, a large inhibitor library may have possible leads for several parasitic diseases (Sajid and McKerrow, 2002).

Cathepsin B is different from cathepsin L in that the pro-domain is about 30 amino acids shorter (Fig. 1.5, panel A) and an occluding loop is present. Furthermore, cathepsin L-like cysteine proteases have a conserved ERFNIN motif which is absent in cathepsin B (Turk *et al.*, 1996; Lecaille *et al.*, 2002). Papain contains two domains: an N-terminal domain with alpha helices and the C-terminal domain with beta sheets, with the catalytic domain located in between (Fig. 1.5, panel B). The active site residues are Cys[25], His[159] and Asn[175] which are conserved among all cysteine proteases (Sajid and McKerrow, 2002)
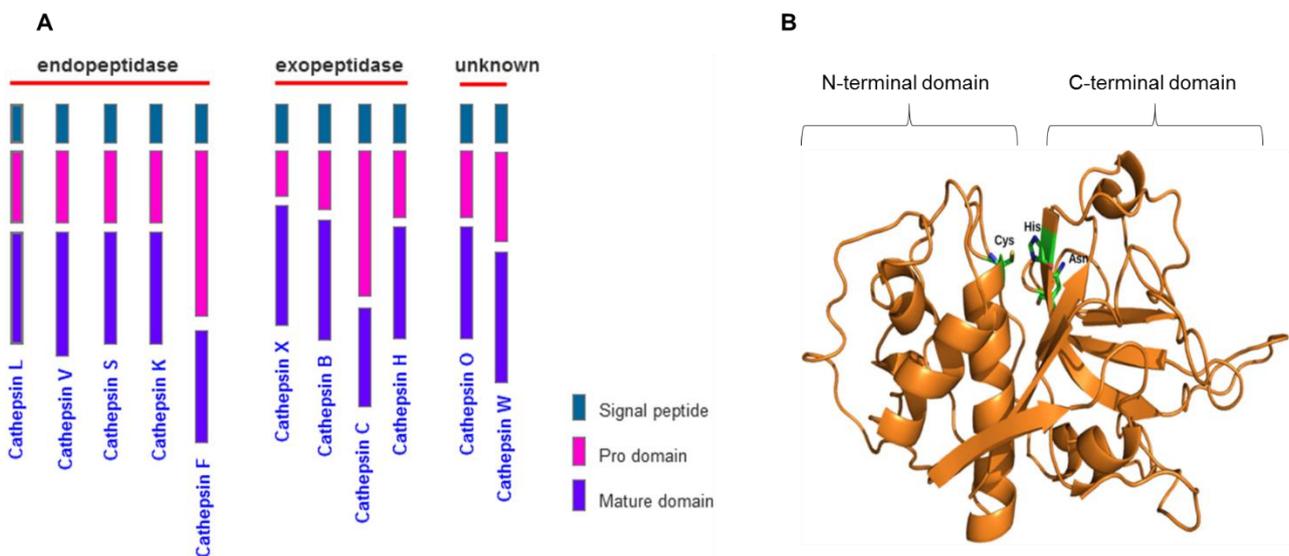
Figure 1.5: Structural representation of the pro and mature domain of cysteine proteases. A, classification of the cysteine proteases and predicted lengths of the signal peptide, pro and mature domain. Accessed from http://www.sinobiological.com/cathepsin-enzyme.html. B, Ribbon diagram showing the N-terminal domain with alpha helices and the C-terminal domain with beta sheets. The catalytic residues Cys, His and Asn are shown as stick models. Adapted from Verma *et al.* (2016).

Papain-like cathepsins are synthesised as inactive precursors (zymogens), to prevent unwanted protein degradation (Verma *et al.*, 2016). They have a pro-domain and a mature (catalytic) domain (Fig. 1.5, panel A) linked by a few Phe and Tyr residues. The pro-domain prevents access to the active site by the substrate, thus acts as an endogenous inhibitor. It also assists in the folding of protein by acting as an intramolecular chaperone. Cysteine proteases may be secreted or targeted to intracellular compartments by an N-terminal signal peptide which comprises of ~15 - 22 hydrophobic amino acid residues (Sajid and McKerrow, 2002).

The interaction between the substrate and the cysteine protease substrate-binding pocket is best described by the Schechter and Berger diagram (Fig. 1.6). This diagram shows that substrate contains amino acid at positions P_1, P_2, P_3 etc. N-terminal to the scissile bond and P'_1, P'_2, P'_3 etc. C-terminal to the scissile bond and the protease contains subsite pockets S_1, S_2, S_3, etc. and S'_1, S'_2, S'_3 etc. which binds amino acids of the substrate in the corresponding positions (Schechter and Berger, 1967). The amino acids found on either side of the scissile bond are important for the specificity of the enzyme. Studies have shown that only five subsites are crucial for substrate binding: S_1, S_2 and S_2' bind the backbone and side-chain of the substrate while S_3 and S_2' bind only the amino acid residue side chains (Turk *et al.*, 2012). Cysteine proteases have substrate specificity at the S_2 pocket and are sensitive to the class

specific cysteine protease inhibitor L-trans-epoxysuccinyl-leucyl- amido (4-guanidino) butane (E-64). Their catalytic mechanism is similar to that of serine proteases, but they have an additional shell of electrons in the thiol group, making them better nucleophiles than the hydroxyl group on serine proteases (Sajid and McKerrow, 2002).
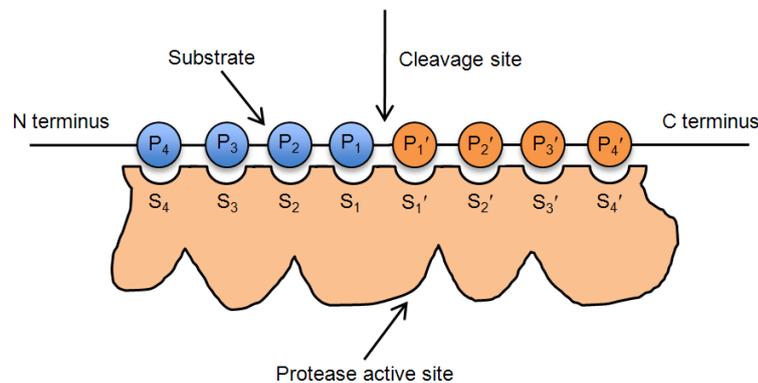
Figure 1.6: Schematic representation of the interaction between enzyme and substrate illustrating the Schechter and Berger (1967) nomenclature. Accessed from https://prosper.erc.monash.edu.au/methodology.html.

1.8.2 Catalytic mechanism of cysteine proteases

Proteases that cleave within the polypeptide are known as endopeptidases while those that cleave at the ends of the polypeptide are known as exopeptidases. Aminopeptidases cleave from the N-terminus and carboxypeptidases cleave from the C-terminus of the substrates. In cysteine proteases, pH changes play a critical role in the activation of the enzyme. Endo-peptidases are activated via auto-activation which occurs in an acidic environment. During this process, the low pH in lysosomes disrupts the interactions between the pro-domain and mature (catalytic) domain, causing them to dissociate to release an active catalytic domain (Turk *et al.*, 2012).

The active enzyme must bind to the substrate prior to hydrolysis (Fig. 1.6). Catalytic residues His, Cys and Asn form the catalytic triad. The His residue acts as a proton donor while the thiol group in Cys acts as a nucleophile, forming an ion pair which is stabilised by Asn. This ion pair attacks the carbonyl group in the scissile bond of the substrate, resulting in a tetrahedral intermediate stabilised by an oxyanion hole formed by asparagine and a highly conserved glutamine residue. The tetrahedral intermediate transforms to an acyl enzyme and the C-terminal end of the scissile bond is cleaved off. The acyl enzyme is then hydrolysed by

water resulting in a second tetrahedral intermediate which ultimately splits into a free enzyme and N-terminal end of the substrate (Verma *et al.*, 2016).
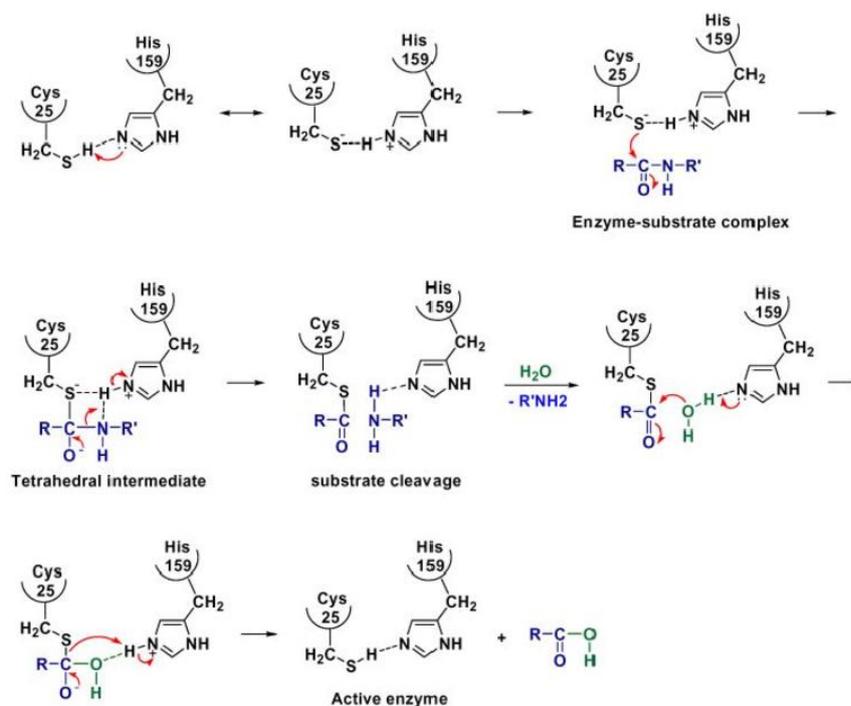
Figure 1. 7: Schematic representation of the catalytic mechanism of cysteine proteases. Accessed from http.researcs://wwwhgate.net/figure/Catalytic-mechanism-of-lysosomal-cysteine-proteinases-as-exemplified-by-papain_fig1_26538859

1.9 Parasite cysteine proteases

Cysteine proteases have been reported in many parasites such as *Plasmodium falciparum*, which causes malaria (Rosenthal *et al.*, 1988), *Fasciola hepatica* which is affects the intestines of mammalian hosts (Chung *et al.*, 1995), *Entamoeba hystolytica* which causes amebiasis (Reed *et al.*, 1993) and *Caenorhabditis elegans* which also affects the intestines of mammalian hosts (McKerrow, 1995). Many of these proteases are papain-like cathepsins B and L and play a major role in pathogenesis. The cysteine proteases of *F. hepatica, E. hystolytica* and *C. elegans* play a role in host tissue invasion and nutrient acquisition (Siqueira-Neto *et al.*, 2018). *P. falciparum* cysteine proteases have been reported to play a role in haemoglobin hydrolysis and red blood cell lysis. Furthermore, cysteine protease inhibitors were found to have antimalarial activity (Marco and Miguel Coteron, 2012). Studies have also shown the potential use of helminth cysteine proteases as drug targets. Mice which were challenged with lung or adult stage *Schistosoma mansoni* worms showed a decrease in worm burden following

treatment with cysteine protease inhibitors (fluoromethyl ketone and vinyl sulfone inhibitors (Caffrey *et al.*, 2018).

1.9.1 Trypanosomal cysteine proteases

Many papain-like cysteine proteases have been reported in *Leishmania* and *Trypanosoma* parasite and the effects of specific inhibitors suggest their role in pathogenesis. Studies on *T. brucei* revealed that cathepsin L protects the parasite from degradation and opsonisation and assists in the crossing of the blood-brain barrier (Nikolskaia *et al.*, 2006). Cathepsin B is found in the BSF of the parasite where it hydrolyses transferrin taken up from the host as part of the iron acquisition pathway (Siqueira-Neto *et al.*, 2018). Other trypanosomal cysteine proteases which have been studied include rhodesain (from *T. b. rhodesiense*), cruzain (from *T. cruzi*), congopain (from *T. congolense*) and vivapain (from *T. vivax*) and were found to be the dominant antigens during infections. RNA interference of *Trypanosoma brucei* cathepsin B and L was also found to reduce the ability of the parasite to cross the blood-brain barrier and prolonged the survival period in mouse models (Abdulla *et al.*, 2008). Vinyl sulfone K777, a potent cruzain inhibitor, was found to decreased parasite burden in animal models challenged with *T. cruzi* (Ferreira and Andricopulo, 2017).

Trypanosomal cathepsin L-like cysteine proteases have a unique C-terminal extension (not found in mammalian homologues) which is highly immunogenic, thereby diverting the immune response away from the functional domain (Boulangé *et al.*, 2001). Congopain (*Tco*CATL) and vivapain (*Tvi*CATL) protease activity have been shown to change during the different life cycle stages of the parasite with the BSF having the highest activity (Mbawa *et al.*, 1992). These proteases are released into the bloodstream after parasite lysis where they contribute to pathogenesis associated with *T. congolense* and *T. vivax* infections. Trypanopain (*Tbb*CATL) was found to be involved in pathogenesis by degrading various host proteins. This enzyme is also released into the bloodstream during *T. b. brucei* infections but unlikely to remain active due to the presence of host inhibitors (Troeberg *et al.*, 1997). Furthermore, vinyl sulfones and peptidyl methyl-ketones which are novel irreversible inhibitors were shown to inhibit the activity of BSF *Tbb*CATL (Troeberg *et al.*, 1999). Consequently, trypanosomal cysteine proteases are thought to play a role in the parasite viability and are targets in chemotherapy and diagnostic developments.

Trypanosomal cysteine proteases produce an antibody response which may not affect the survival of the parasite, however, they are believed to play a role in trypanotolerance by neutralising the effects of the enzyme. Therefore, the effects of antibodies on the enzyme activity have been investigated. Troeberg *et al.* (1997) reported the presence of both activity-enhancing and activity-inhibiting antibody responses to *Tbb*CATL due to the polyclonal nature

of the antibodies. On the other hand, anti-congopain IgG as well as antigen binding fragments (Fab) showed inhibitory effects on congopain activity against the peptide substrate Z-Phe-Arg-AMC which indicate that host-specific antibodies may mitigate trypanosome pathology (Authié *et al.*, 2001; Lalmanach *et al.*, 2002).

It is evident from the findings that cysteine proteases are desirable targets in the development of antigen based diagnostic tools. Although antigen-based diagnostic kits are an indication of current infection, the sensitivity of the tests depends on the number of circulating parasites antigen in the host. Furthermore, the target antigens must have epitopes specific to each *Trypanosoma* spp. Therefore, there is an additional requirement for capture antibodies which recognise the distinct epitopes on each antigen and do not interact with the host IgG antibodies resulting in false positives. This is often occurring from interactions with the constant Fc regions of the antibodies (Torres *et al.*, 2018). This can be overcome by using substitute antibody molecules recruited from large libraries of peptides and functional domains of different proteins displayed on the surface of bacteriophages. This technology is known as phage display technology (discussed in Section 1.9) (Smith and Scott, 1993).

1.10 Phage display technology

Phage display technology is a method used to display polypeptides or proteins on the surface of filamentous bacteriophages (viruses that infect Gram negative bacteria). These phages are said to be "friendly" phages since they can assemble and replicate without killing the host cell. The polypeptide libraries presented on the surface of phages allow the selection of polypeptides, proteins and functional antibody fragments with high affinity and specificity to a target of interest (Arap, 2005). They are usually displayed as fusions with the phage coat proteins. This is achieved by fusing genes encoding polypeptides or proteins with genes coding the phage coat proteins thus allowing the simultaneous construction of large phage display libraries (up to 10^{10} variants) (Arap, 2005; Paschke, 2006).

1.10.1 General biology of filamentous phages

Bacteriophage species used in phage display technology include f1, fd, T4, M13 and T7. These phages are commonly used due to their potential as excellent cloning vehicles and their ability to correctly assemble longer phage particles in a simple manner. Phages have single stranded DNA (ssDNA) which encodes three classes of proteins: replication proteins, morphogenic proteins and structural proteins. Structural proteins make up a large component of the phage particles and are encoded by *pIII, pVI, pVII, pVIII* and *pIX* genes of the single stranded genome (Fig. 1.8). The *pVIII* gene encodes the major coat protein which makes up 99% of the protein mass and the remainder is composed of the minor coat protein encoded

by *pIII, pVI, pVII* and *pIX* (Endemann and Model, 1995). The coat proteins act as anchors for the displayed polypeptide and do not interfere with peptide-ligand interaction (Paschke, 2006).
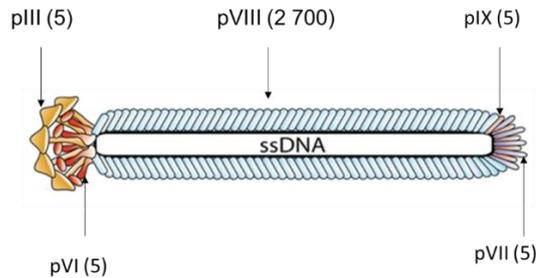
Figure 1.8: **Schematic representation of the filamentous phage M13 and coat proteins.** The single stranded phage is encapsulated by 2 700 copies of the major coat proteins, pVIII and 5 copies of each of the minor coat proteins pIII, pVI, pIX and pVII (Løset *et al.*, 2011).

The most commonly used phages, such as the filamentous phage M13, use minor protein pIII and major protein pVIII as fusion proteins (Fig. 1.8) (Smith and Scott, 1993; Endemann and Model, 1995). The pIII library can display up to five copies of the display peptides while the pVIII can display up to 2700 copies of small peptides. Additionally, M13 phages can only display peptides and proteins that are able to pass through the bacterial membrane while maintaining the proper folding (Scott and Smith, 1990; Smith and Scott, 1993).

1.10.2 Vectors in phage display technology

In early studies, the genes of an antibody fragment or polypeptide were cloned directly into the phage genome. Since the phage displays three to five copies of the pIII minor coat protein, phage vectors often resulted in the display of multivalent copies of the antibody fragment or polypeptide, thereby decreasing the infectivity of the phage. Additionally, it is difficult to achieve high affinity selection with multivalent display vectors because they do not distinguish between weak and strong affinity ligands (Petrenko and Vodyanoy, 2003). An alternative approach was the use of phagemid vectors. Phagemid vectors are designed by combining the phage genome and portions of the plasmid. Like the phage vector, they contain essential structural components i.e. a promoter region, signal peptide, cloning site, detection and purification marker (tag), amber codon for expression, coat protein gene, origin of replication for filamentous phage, *Escherichia coli* replication origin and selection marker (antibiotic resistance gene) (Fig. 1.9). Phagemid vectors are mostly used because they have higher transformation efficiencies compared to phage vectors and allow for the construction of larger libraries.

The size of the expressed foreign protein limits the use of the phage. Although the pIII protein can be used to display large peptides, the infectivity of the phage is compromised. To overcome this, bacteria containing phagemid vectors are infected with a helper phage such as the M13KO7 helper phage. The helper phage contains all the genes required for packaging, infection, replication, assembly and budding of the phagemid, but has a low copy origin of replication which causes the helper phage to be replicated but packaged less effectively (Winter *et al.*, 1994). The modified origin of replication is inserted into the native M13 origin of replication. During infection, the wildtype pIII from the helper phage competes with the pIII-fusion protein from the phagemid for expression on the surface of the phage. The majority of the phages express the wildtype pIII protein (known as the "bald phages") while only a few express the pIII-fusion protein.
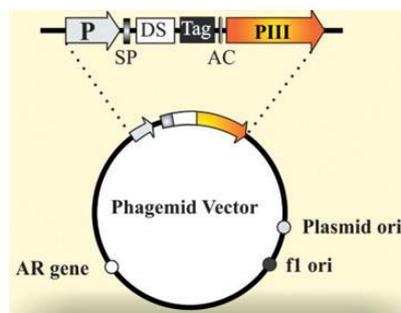
Figure 1.9: Schematic representation of construction of phagemid vector for phage display library. The vector contains an origin of replication of ssDNA (f1 ori), plasmid origin of replication for *E. coli* and antibiotic resistance gene (AR). It has a promotor (P), the display sequence (DS) cloned between the signal peptide (SP) and pIII coat gene. The amber codon (AC) is located between the DS and the pIII (Zhao *et al.*, 2016).

During protein synthesis, the amber codon recruits a tRNA carrying a complementary stop codon which signals termination of proteins synthesis. Certain cells produce a mutation in the DNA complementary to the amber codon, preventing the signal to terminate protein synthesis. Such cells are known as suppressors (Cropp and Schultz, 2004). In phagemids, the amber codon of protein synthesis is often inserted between the gene for pIII and the cloned fragment (Fig. 1.9). This allows the fragment and fusion protein to be expressed on the surface in suppressor cells, or expressed in the soluble form in non-suppressor *E. coli* cells (Hoogenboom *et al.*, 1991).

The pIII protein is located on the surface of the phage capsid and contains three functional autonomous domains i.e. D1 (at the C-terminus), D2 and D3 (at the N-terminus). During infection of *E. coli*, D1 translocates the viral DNA into the host cytoplasm, D2 binds to the bacterial F-pilus and D3 plays a role in capsid assembly and reproduction (Georgieva and

Konthur, 2011). The advantage of using pIII is that is allows incorporation and presentation of larger polypeptides compared to pVIII. Infection begins by binding of the phage to the pilus of the bacterium, followed by conversion of the ssDNA to dsDNA (replicative form) by cytoplasmic enzymes. The phage particles are assembled from the inner membrane of the bacterium and are secreted with the help of pVII and pIX (Rami *et al.*, 2017).

1.10.3 Phage display technology in antibody scFv fragment production

Antibodies can bind a large variety of antigens with high affinity and specificity, and have important properties required for diagnostic development, however, due to their complexity, they cannot be produced in large amounts or functional forms in bacterial hosts. Additionally, the size of the displayed antibody may reduce the infectivity of the phage previously mentioned (Section 1.10.2) (Ruigrok *et al.*, 2011). The immunoglobulin G (IgG) is a heterodimer which consists of two identical heavy and light chains linked by disulfide bonds (Fig. 1.10). Recombinant expression hosts do not have optimal conditions required for the expression of a complete antibody molecule therefore, antibody fragments are used in *E. coli* -based phage display systems.

Different antibody fragments can be displayed on the surface of bacteriophages, these include the antigen binding fragment (Fab), variable fragment (Fv) and single chain variable fragments (scFv). The Fv is the smallest immunoglobulin fragment with antigen binding activity. Single chain variable fragment (scFv) consists of two Fv fragments of the heavy chain and light chain linked by a short peptide linker of about 25 amino acids to improve the stability of the scFv. The peptide linker consists of glycine, and serine or threonine amino acid residues, making it flexible and soluble (Kuhn *et al.*, 2016). Despite the absence of the constant regions and addition of a peptide linker in scFvs, the specificity of the original immunoglobulin remains intact. Additionally, being half the size of the Fab with the same specificity as the parent immunoglobulin, scFvs are easy to express and have been successfully applied in techniques such as flow cytometry and immuno-histochemistry (Arap, 2005).
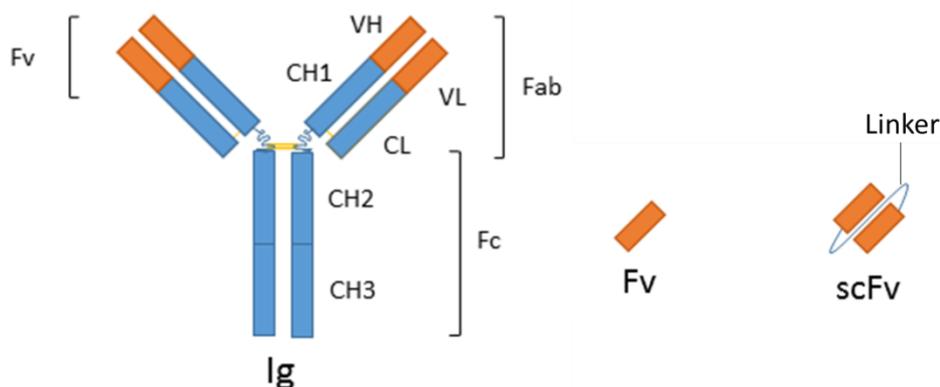
Figure 1.10: Schematic representation of the antibody and derived fragments. (Ig): immunoglobulin made up of the heavy chain (Y shaped), light chain, variable fragment (Fv); Fab (fragment antigen binding); Fc (fragment crystallisable), CH1-3 (constant regions of the heavy chain); CL (constant region of the light chain); scFv (single chain variable fragments) joined by a linker. Accessed from https://www.antibodies-online.com/resources/18/1502/antibody-and-immunoglobulin-alternatives-part-1/

1.10.4 Phage display antibody libraries

There are four main types of antibody phage libraries i.e. immune, naive, semi-synthetic and synthetic libraries. In immune libraries, the B-cells are isolated from the spleen of an immunised animal, the V genes from the IgG mRNA are cloned into the phage library vectors. The advantage of immune libraries is that the antibodies isolated have a high affinity to their specific antigen because affinity maturation has already occurred in the immune system of the immunised animal. This technique allows antibodies to be isolated and rapidly produced from a single immunised donor and further manipulated. Immune libraries have been produced from various species such as humans (Hoogenboom and Winter, 1992), chickens (Sixholo *et al.*, 2011) and rabbits (Weber *et al.*, 2017), however, due to ethical concerns, active immunisations are only possible in some experimental animals. Some animals may have tolerance towards the antigen making the procedure less effective. Furthermore, each antigen requires the construction of a new library which can be time consuming and costly, therefore, different types of non-immune libraries were introduced to overcome these limitations (Iba and Kurosawa, 1997).

The non-immune libraries, also known as the "single pot'' library, include the naive, synthetic and semi-synthetic libraries. They are constructed from B-cells isolated from the spleen of animals. In a naïve library, the V genes are derived from the IgM mRNA of a non-immunised animal (Rami *et al.*, 2017). The genes are amplified from the cDNA followed by random combination of the V_H and V_L genes which are cloned to form scFvs or a Fab. This technique

produces antibodies which have not been exposed to antigens, but can bind to different antigens including self, non-immunogenic as well as toxic antigens (Hoogenboom *et al.*, 1998).

Synthetic libraries are constructed by introducing diversity in the complementary determining region (CDR) of the hypervariable regions (Rami *et al.*, 2017). The diversity can be manipulated to correspond to areas where diversity occurs naturally in the antibody repertoire. Most of the natural sequence and structural diversity occurs in CDR3 of the heavy chain, making it a target for the introduction of diversity, while the other five CDRs have limited diversity (Conrad and Scheller, 2005). This library has an advantage over naïve libraries, where rearrangement of the V genes occurs naturally because the antibody repertoires can be constructed in a manner that allows optimum performance of the library. Semi-synthetic libraries are constructed by substituting the amino acids of the CDR regions which are involved in the binding of antigen to produce a library with high diversity that does not exist in vivo (Zhao *et al.*, 2016).

The most successful use of these libraries is in the screening and selection of monoclonal antibody fragments from a large library (Winter *et al.*, 1994; Ahmad *et al.*, 2012). To achieve this, the phages are constructed and released from the host in large amounts to ensure the likelihood of the desired fragment being present. This is determined by the binding capacity of the fragment to a specific ligand (Rakonjac, 2012).

1.10.5 Biopanning

Once the V genes are cloned into phagemid vectors, a variety of scFvs are displayed on the surface of the phage. These phages are then allowed to interact with an antigen of interest to select for specific clones with high affinity for the antigen in a process known as biopanning (or panning for short) (Petrenko and Vodyanoy, 2003). The high affinity antibodies are enriched by several rounds of panning, normally three to five rounds. There are many procedures that have been described for the panning of phages but the most common is done on the immobilised antigen (Hoogenboom *et al.*, 1998).

The panning process includes binding of the antigen onto a solid surface, followed by a blocking step, incubation of phage with the immobilised antigen followed by a washing step to remove non-specific phages. The phages are eluted and infected in bacterial cells for amplification of eluted phages (Fig.1.11). The bound phages are eluted with either an acidic buffer or with a basic buffer (Ruigrok *et al.*, 2011). Other methods of elution include the use of dithiothreitol, enzymatic cleavage at the cleavage site inserted between the antigen and the

phage or addition of excess antigen to outcompete the bound phages (Carmen and Jermutus, 2002).
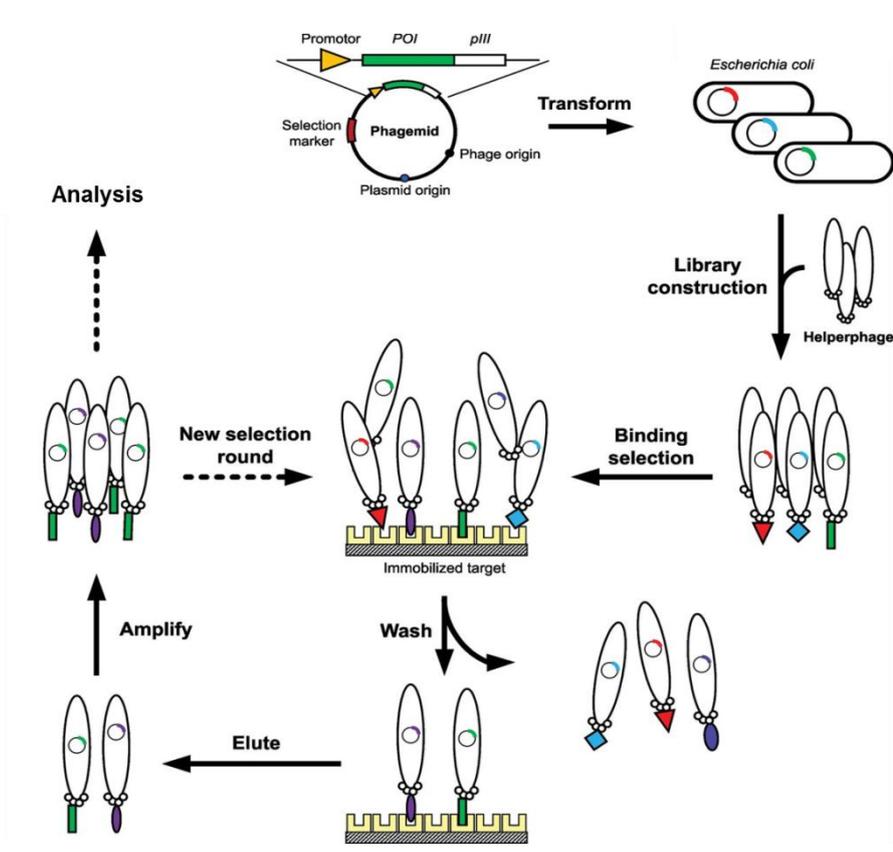
Figure 1.11: Phage cycle and biopanning (Ruigrok *et al.*, 2011)

During the first round of panning, the number of phages that bind is far less than the number that does not bind. Therefore, this round requires a high enough concentration of the antigen to increase the chances of selecting specific clones which may be present in low concentrations. These clones are amplified in *E. coli* and used for subsequent panning rounds (Carmen and Jermutus, 2002). After panning, the phages are then screened using immunoassays or immunocytochemistry with the purpose of isolating those with high affinity to the target antigen (Ruigrok *et al.*, 2011).

1.11 Aims and objectives of the study

1.11.1 General objective

Cysteine proteases from *T. congolense*, *T. vivax* and *T. b. brucei* have been identified as potential diagnostic targets for AAT. The main objective of this study was to produce antibodies which can be used for antigen-based specific diagnosis of *T. b. brucei* infection, targeting the cysteine protease (*Tbb*CATL). Previous studies of *Tbb*CATL were conducted on the native form of the enzyme and there are no reports on the recombinant form. The homologues from *T. congolense* and *T. vivax* have been recombinantly expressed and purified in their truncated form (excluding the C-terminal extension).

1.11.2 Aims

1. To clone and explore recombinantly expressing *Tbb*CATL in bacteria to avoid possible intellectual property restrictions of the *Pichia pastoris* expression vectors. In addition, *Tco*CATL and *Tvi*CATL were recombinantly expressed in *P. pastoris.*

2. To produce anti-*Tbb*CATL antibodies in chickens and analyse their cross-reactivity with, *Tco*CATL and *Tvi*CATL in a western blot.

3. To optimise phage display technology as an alternative to animal-based antibody production. The Nkuku® phage library was used to select for scFvs specific for *Tvi*CATL for potential use in a rapid diagnostic test.

Chapter 2

Cloning, recombinant expression and purification of trypanosomal cysteine proteases from *T. b. brucei* (*Tbb*CATL), *T. congolense* (*Tco*CATL) and *T. vivax* (*Tvi*CATL)

2.1 Introduction

Trypanosoma congolense, *T. vivax* and *T. b. brucei* cause animal African trypanosomiasis (AAT) or nagana which is an economic burden and hampers agricultural development in Africa (Muhanguzi *et al.*, 2017). The parasite escapes the host's immune response by switching genes coding for the variable surface glycoproteins, resulting in new variable antigen types. This has made it unlikely to develop a vaccine against the disease and therefore many studies now focus on identifying trypanosome antigens as potential drug targets (Giordani *et al.*, 2016). Studies on *T. congolense* revealed that the cathepsin L-like cysteine protease, *Tco*CATL, is involved in the pathogenesis by degrading host proteins (Serveau *et al.*, 2003). In addition, trypanotolerant N'Dama cattle were found to have high levels of IgG against *Tco*CATL in their blood serum. These antibodies were believed to contribute to trypanotolerance by inhibiting *Tco*CATL activity, thus this enzyme is considered a possible candidate for an anti-disease vaccine to limit the symptoms of the disease, such as anaemia, rather than killing the parasite (Authié *et al.*, 2001).

Trypanosoma brucei brucei cathepsin L-like cysteine protease *(TbbCATL)* shares sequence identity with *Tco*CATL and homologues from *T. vivax* (*Tvi*CATL) as well as *Tcr*CATL from *T. cruzi*, the parasite that causes Chagas disease in humans in South America. This suggests that these cathepsins may be similarly involved in pathogenesis and are sought as potential drug targets. Like other clan CA cysteine proteases, these proteases contain a pro-domain, catalytic domain and a unique C-terminal extension which does not occur in mammalian cysteine proteases (Boulangé *et al.*, 2011). The C-terminal extension, whose function is not yet understood, is linked to the catalytic domain by a proline rich hinge in the case of *Tco*CATL and *Tbb*CATL, and a threonine rich hinge in *Tvi*CATL (Fig. 2.1). The catalytic residues Cys^{25} and His^{191} (*Tco*CATL numbering) are spatially positioned to form an ion pair which is stabilised by Asn^{211} during substrate hydrolysis (Sajid and McKerrow, 2002).

Since the sequencing of the cysteine protease *Tbb*CATL from *T. b. brucei* cDNA (Mottram *et al.*, 1989), this protease has been isolated, purified and characterised through substrate specificity and inhibition studies (Troeberg *et al.*, 1999). These studies were done on the native

protease and to date, there are no reports on recombinantly expressed *Tbb*CATL. Although recombinant expression of an active full length and truncated *Tcr*CATL was previously reported in *E. coli* (Eakin *et al.*, 1993), recombinant expression of *Tco*CATL has been confined to the 28 kDa catalytic domain excluding the C-terminal extension as it is highly immunogenic (Authié *et al.*, 2001). This expression was done in the methylotrophic yeast, *Pichia pastoris* which was reported to yield more *Tco*CATL compared to bacterial expression systems. In light of this, *Tvi*CATL, excluding the C-terminal extension, has also been recombinantly expressed in *P. pastoris* and both proteases have been enzymatically characterised (Boulangé *et al.*, 2011; Eyssen *et al.*, 2018).

There are many advantages of using *P. pastoris,* however, the AOX 1 requires methanol for induction which can be toxic to the cells and is flammable. Yeast cells require a longer biomass generation time prior to the induction phase and switching from glucose as a carbon source in the growth phase to methanol in the induction phase at the exact time is difficult. Additionally, there are possible intellectual property restrictions on the *P. pastoris* vector if the produced antigens are to be used for diagnostic assays (Juturu and Wu, 2018). Therefore, the *E. coli* expression is usually the first expression system to be explored because it is quick and inexpensive to culture and high protein yields can be obtained at small scale due to the strong promoter. Competent *E. coli* cells can be stored and thawed before use whereas yeast cells must be used immediately and requires screening of many colonies to select a good producer as protein yield can vary among clones. Furthermore, in *E. coli* proteins are expressed with a purification tag that can also be used to detect the expressed protein using commercial antibodies (Bill, 2014).

Cloning vectors that allow C-terminal or N-terminal positioning of fusion tags are available. Fusion tags can either be peptide tags which include poly-Arg, poly-His, FLAG, or non-peptide tags which include maltose-binding proteins (MBP) and glutathione S-transferase (GST) (Terpe, 2003). These tags are useful in the detection of recombinantly expressed proteins by western blots, especially when the expressed protein is too small to be identified by SDS-PAGE analysis.

In this chapter, *Tbb*CATL was cloned, recombinantly expressed in *E. coli* BL21 (DE3) using the pGEX-4T vector which adds a 26 kDa GST tag and purified. pET32a and pET28a were also used for expression and the results are reported in this chapter. Recombinant expression and purification of *Tco*CATL and *Tvi*CATL in *P. pastoris* is also reported. Further details of the diagnostic target are reported in chapter 3.

```
TcrCATL    -----MSGWARFVLLAAVLVVMACLVPAATASLHAEETLTSQFAEFKQKHGRVYESAAR-   54
TviCATL    -------MHAH--ALVTLLAAAVSVAPAAMAVLRADGPVEPLFAAFKQKYGRSYGTAAEE   51
TbbCATL    MPRTEMVRFVR--LPVVLLAMAACLASVALGSLHVEESLEMRFAAFKKKYGKVYKDAKEE   58
TcoCATL    MPRSEMTRTLR--FSVGLLAVAACFVPVALGVLHAEQSLQQQFAAFKQKYSRSYKDATEE   58
             .  :*.   . :*.  .... .* . *:.:  :    ** **:*:.: *   * .

TcrCATL    RLPLSVFRENLFLARLHAAANPHATFGVTPFSDLTREEFRSRYHNGAAHFAAAQERARVP   114
TviCATL    AFRLRVFEDNMRRSRMYAAANPHATFGVTPFSDLTPEEFRTRYHNGERHFEAARGRVRTL   111
TbbCATL    AFRFRAFEENMEQAKIQAAANPYATFGVTPFSDMTREEFRARYRNGASYFAAAQKRLRKT   118
TcoCATL    AFRFRVFKQNMERAKEEAAANPYATFGVTRFSDMSPEEFRATYHNGAEYYAAALKRPRKV   118
            : : .*.:*:  ::  *****:****** ***:: ****: *:.**  :: **   * *

TcrCATL    VKVEVVGAPAAVDWRARGAVTAVKDQGQCGSCWAFSAIGNVECQWFLAGHPLTNLSEQML   174
TviCATL    VQVPPGKAPAAVDWRRKGAVTPVKDQGTCGSCWSFSAIGNIEGQWAAAGNPLTSLSEQML   171
TbbCATL    VNVTTGRAPAAVDWREKGAVTPVKVQGQCGSCWAFSTIGNIEGQWQVAGNPLVSLSEQML   178
TcoCATL    VNVSTGKAPEAVDWRKKGAVTPVKDQGQCGSCWAFSAIGNIEGQWKVAGHELTSLSEQML   178
            *.*   ** ***** :**** ** ** *****:**:***:* **   **: *..******

TcrCATL    VSCDKTDFGCSGGLMNNAFEWIVQENNGAVYTEDSYPYASGEGISPPCTTSGHTVGATIT   234
TviCATL    VSCDTKDNGCGGGLMDNAFEWIVKENSGKVYTEKSYPYVSGGGEEPPCKPRGHKVGATIT   231
TbbCATL    VSCDTIDSGCNGGLMDNAFNWIVNSNGGNVFTEASYPYVSGNGEQPQCQMNGHEIGAAIT   238
TcoCATL    VSCDTNDFGCEGGLMDDAFKWIVSSNKGNVFTEQSYPYASGGGNVPTCDKSGKVVGAKIR   238
            ****.  * ** ****::**:***..* * *:** ****.** *  * *   *: :** *

TcrCATL    GHVELPQDEAQIAACVAVNGPVAVAVDASSWMTYTGGVMTSCVSEQLDHGVLLVGYNDSA   294
TviCATL    GHVDIPHDEDAIAKYLADNGPVAVAVDATTFMSYSGGVVTSCTSEALNHGVLLVGYNDSS   291
TbbCATL    DHVDLPQDEDAIAAYLAENGPLAIAVDAESFMDYNGGILTSCTSKQLDHGVLLVGYNDNS   298
TcoCATL    DHVDLPEDENAIAEWLAKNGPVAIAVDATSFQSYTGGVLTSCISEHLDHGVLLVGYDDTS   298
            .**::*.** **  :* ***:*:**** ::  *.**::*** *: *:*********:*.:

TcrCATL    AVPYWIIKNSWTT-QGEEGYIRIAKGSNQCLVKEEASSAVVGGPGPTPEPTTTTTTSAPG   353
TviCATL    KPPYWIIKNSWSSSWGEKGYIRIEKGTNQCLVAQLASSAVVGGPGPTPTPTPT-------   344
TbbCATL    NPPYWIIKNSWSNMWGEDGYIRIEKGTNQCLMNQAVSSAVVGGPTPPPPPPPP-------   351
TcoCATL    KPPYWIIKNSWSKGWEEGYSALRR-HNQCLMKNLPSSAVVSGPP-----PPP-------   345
            *********:. **.** :: ****: :  *****.**

TcrCATL    PSPSYFVQMSCTDAACIVGCENVTLPTGQCLLTTSGVSAIVTCGAETLTEEVFLTSTHCS   413
TviCATL    -TNNNNNND-------------RTWPIVKLHEDA------------LQR----------   367
TbbCATL    -PSATFTQDFCEGKGCTKGCSHATFPTGECVQTTGVGSVIATCGASNLTQIIYPLSRSCS   410
TcoCATL    -PTPTFTQELCEGAECQSKCTKATFPTGKCVQLSGAGSVIASCGSNNLTQIVYPLSSSCS   404
                :             * * :   :           * .

TcrCATL    GPSVRSSVPLNKCNRLLRGSVEFFCGSSSSGRLATWTGSVAISHTTAVIAASEGTLCF   471
TviCATL    ---------------------------------------------------------   367
TbbCATL    GPSVPITVPLDKCIPILIGSVEYHCSTNPPTKAARLVPHQ-----------------   450
TcoCATL    GFSVPLTVPLDKCLPIVIGSVMYECSDKAPTESARLVRHE-----------------   444
```

Figure 2.1: Sequence alignments of cathepsin L-like cysteine proteases from *T. b. brucei* (*Tbb*CATL), *T. congolense* (*Tco*CATL), *T. vivax* (*Tvi*CATL) and *T. cruzi* (*Tcr*CATL). The sequences were aligned using Clustal Omega multiple sequence alignment tool. *Tbb*CATL accession number CAA34485.1, *Tco*CATL accession number AAA18215.1, *Tvi*CATL accession number CCD21670.1, *Tcr*CATL accession number AACC00067.1. The N-terminal pro-domain sequence (11-12 kDa) is enclosed in the green shape, catalytic domain sequence (28-31 kDa) is enclosed in the red shape and the C-terminal extension sequence (11-13 kDa) is enclosed in the purple shape. The catalytic residues C, H and N are indicated by the arrows. The predicted Asn glycosylation site for *T. vivax* is underlined within the catalytic domain. The peptide sequences highlighted in blue were used in antibody production (discussed in Chapter 3).

2.2 Materials and methods

2.2.1 Materials

Molecular Biology: High fidelity PCR enzyme mix, GeneJet™ Plasmid Miniprep kit, O'geneRuler™ 1Kb DNA Ladder, 5-bromo-4-chloro-3-indolyl-β-D-galactopyronoside (IPTG), isopropyl-β-D-thiogalactopyronoside (X-gal), T4 DNA ligase and PTZ57R/T were purchased from Themo Scientific (USA). Zymo Research DNA clean and concentrator™5 kit was purchased from ZymoResearch (USA). BamHI, XhoI, InsTAclone PCR cloning kit and TransformAid™ Bacterial Transformation kit were purchased from Fermentas (Vilnius, Lithuania). The E.Z.N.A® gel extraction kit was from PEQlab biotechnology (Germany), pGEM®_T cloning vector was purchased from Promega (Madison, WI, USA) and pET32a and pET28a were purchased from Novagen (Darmstadt, Germany). pGEXT-4T was purchased from GE Healthcare (Uppsala, Sweden). FIREpol® Taq polymerase, 10 x PCR reaction buffer and 25 mM $MgCl_2$ were obtained from Solis Biodyne (Tartu, Estonia). Crystal violet and ethidium bromide were from Sigma-Aldrich (USA). Seakem® LE agarose was purchased from Lonza (USA) and tryptone, yeast extract and bacteriological agar were purchased from Merck (Darmstadt, Germany). Ampicillin was purchased from Amresco (USA) and kanamycin from Life Technologies (UK).

***Escherichia coli* cells:** *E. coli* strains JM109 and BL21 (DE3) were purchased from New England Biolabs (Ipswitch, MA, USA). *E. coli* JM109 allows blue/white colony screening of transformants in the presence of IPTG and X-gal, while *E. coli* BL21 (DE3) lacks Lon and OmpT protease expression (Miroux and Walker, 1996).

Antibodies: Mouse anti-6xHis IgG-horseradish peroxidase (HRPO) and rabbit anti-chicken IgY-HRPO detection antibodies were from Sigma-Aldrich (USA) and chicken anti- glutathione-S-transferase (GST) IgY antibodies were produced in-house.

Glycerol stocks of plasmid containing *Tvi*CATL and *Tco*CATL genes: Glycerol stocks of *P. pastoris* GS115 containing *Tvi*CATL was obtained from Professor Theo Baltz (University of Victor Segalen, Bordeaux, France) and *Tco*CATL was from Dr Alain Boulangé (University of KwaZulu-Natal, Pietermaritzburg). Tetracycline, yeast nitrogen base (without amino acids) and biotin were purchased from Sigma-Aldrich (USA).

Protein purification and quantification: Glutathione-agarose resin was purchased from Sigma-Aldrich (USA). The HiPrep® Sephacryl® S-200 HR prepacked column and ÄKTA prime plus were purchased from GE Healthcare (Uppsala, Sweden). Tertiary butanol and ammonium sulfate were from Sigma-Aldrich (USA). Phosphorylase B (97.4 kDa), bovine

serum albumin (66 kDa), ovalbumin (45 kDa) and carbonic anhydrase (30 kDa) were purchased from Sigma-Aldrich (USA).

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and western blot analysis: Hydrogen peroxide (H_2O_2) and 2-mercaptoethanol were purchased from Merck (RSA). PageRuler pre-stained marker was purchased from Separation Scientific (SA), BioTrace™ nitrocellulose was from PALL Life Sciences (USA) and 4-chloro-1-naphthol was from ThermoScientific (Walthem, USA).

Peptide substrates and inhibitors: The peptide substrate Z-Phe-Arg-amino-4-methylcoumarin (AMC), H-D-Val-Leu-Lys-AMC and L-trans-epoxysuccinyl-leucylamido (4-guanidino) butane (E-64) were purchased from Sigma-Aldrich (USA).

2.2.2 Bradford Assay for protein quantification

The Bradford dye-binding assay (Bradford, 1976) was done to determine the concentration of the purified proteins. A 1 mg/ml stock solution of ovalbumin was used to construct a standard curve (Fig. 2.2). Triplicate samples of 20-100 µg/ml in 100 µl dH_2O were prepared, 900 µl of the Bradford dye reagent [0.6% (w/v) Coomassie brilliant blue G-250 in 2% (v/v) perchloric acid] was added and mixed by vortexing and incubated for 5 min at RT. The absorbance was measured at 595 nm.
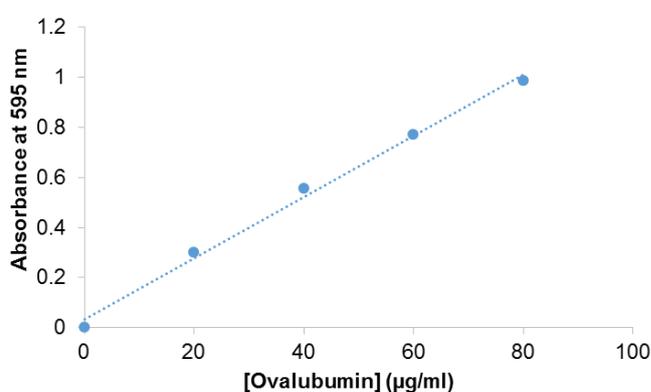
Figure 2.2*: Bradford dye-binding assay standard curve.* The standard curve was constructed from ovalbumin standards ranging from 20-100 µg/ml and allowed to interact with the Bradford dye reagent [0.6% (w/v) Coomassie brilliant blue G-250 in 2% (v/v) perchloric acid]. The absorbance was measured at 595 nm. The trend line equation was y=0.0122x + 0.0346 with a correlation coefficient of 0.99.

2.2.3 Cloning of *TbbCATL* into pGEM®-T and PTZ57R/T TA cloning vectors

The primers for amplification of the gene coding for the catalytic domain of *T. b. brucei* cathepsin L-like protease were: forward primer: 5'-GC GGATCC ATG GCG TGC CTT GCG

TCT GTC G-3' and <u>reverse primer</u>: 5'- GC <u>CTCGAG</u> TCA ACT GCG GAG ACG GCC TG -3'. The restriction site for BamHI (forward) and XhoI (reverse) are underlined.

Genomic DNA of *T. brucei* available in our laboratory was used as a template for the amplification of the *TbbCATL* gene. The PCR master mix contained *Tbb*CATL Fw and Rv primers (0.25 µM for each), 1x High Fidelity PCR enzyme mix buffer, 2.5 mM $MgCl_2$, 0.25 U High Fidelity Taq polymerase and 0.5 mM dNTPs in a 20 µl reaction volume for 35 cycles with the conditions listed in Table 2.1. A sample of the PCR product (5 µl) was analysed by gel electrophoresis (Section 2.2.4). The remaining sample was purified using Zymo Research DNA gel recovery kit according to the manufacturers' instruction manual. To determine which cloning vector works best, the purified *TbbCATL* gene fragment was ligated into pGEM®-T and PT757R/T cloning vectors (Fig. 2.3) in a 3:1 ratio of gene insert to vector using 6 U of T4 DNA ligase and incubated at 37°C for 1 hour followed by RT for 16 hours.
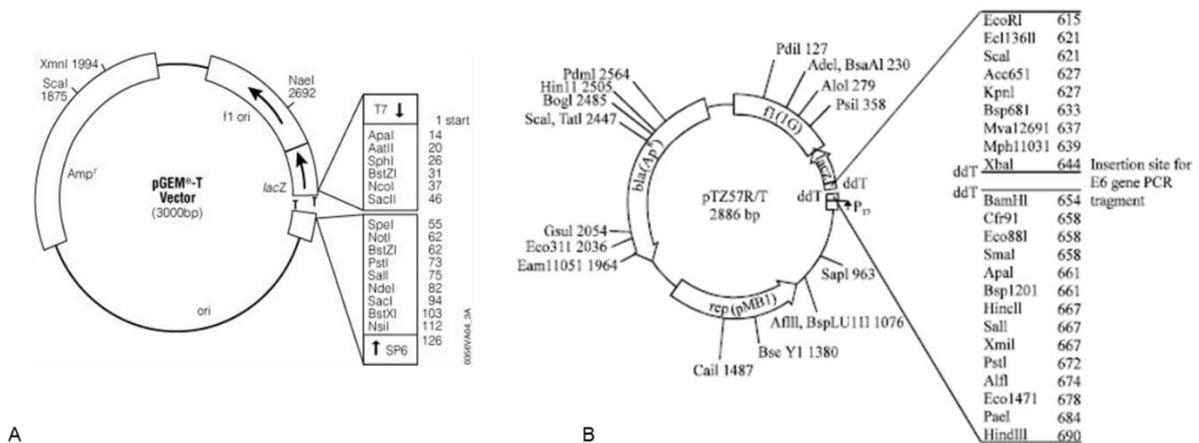
Figure 2.3: TA cloning vector maps. (A) pGEM®-T (Promega); (B) PTZ57R/T (Thermo Fisher Scientific)

Table 2.1: PCR conditions for amplification of the *TbbCATL* gene from genomic DNA

Cycle step	Temperature (⁰C)	Duration
Initial denaturation	94	3 min
Denaturation	94	30 s
Annealing	53	30 s
Extension	72	1 min
Final extension	72	10 min

2.2.4 Gel electrophoresis for DNA analysis

A sample of the PCR product (5 µl) was electrophoresed on a 1% (w/v) agarose gel containing 0.5 µg/ml ethidium bromide in Tris-Acetate-EDTA (TAE) buffer (40 mM Tris-HCl, 100 mM acetic acid, 1 mM Na$_2$EDTA, pH 8.0). The image was captured using the G:Box system (Syngene, Cambridge UK). A standard curve for the molecular weight marker separated on a 1% agarose gel was constructed (Fig. 2.4).
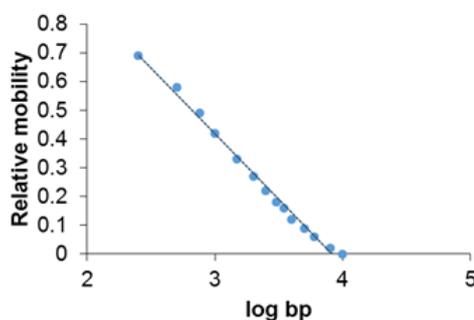
Figure 2.4: Standard curve for relative mobility against log of base pairs used in agarose 1% gels. Commercial O'geneRuler™ 1kb DNA Ladder was used in agarose gel electrophoresis. The trendline equation was given by y=-0.04579x+1.7916 with a correlation coefficient of 0.99.

2.2.5 Transformation of recombinant *TbbCATL*_pGEM®-T and *TbbCATL*_PTZ57R/T into *E. coli* JM109 cells and colony PCR

The ligation mixture was transformed into *E. coli* JM109 cells using the TransformAid™ bacterial transformation kit according to the manufacturer's instruction manual, plated onto pre-warmed 2xYT medium [1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl, containing 1% (w/v) bacteriological agar, 20 µg/ml X-gal, 50 µg/ml ampicillin and 1 mM IPTG] and incubated at 37⁰C overnight. A mixture of blue and white colonies was obtained from which a white colony was randomly selected and resuspended in 10 µl of sterilised distilled water. The colony mixture (2 µl) was used in colony PCR which was carried out (as per Section 2.2.3, with 5 U FIREPol® Taq polymerase) in a 50 µl reaction volume for 25 cycles using the conditions shown in Table 2.2. A sample of the PCR product was electrophoresed on a 1% (w/v) agarose gel containing 0.5 µg/ml ethidium bromide. From the remaining colony mixture, 2 µl was inoculated in 6 ml 2xYT broth medium containing 50 µg/ml ampicillin and incubated at 37⁰C overnight, with agitation (200 rpm).

Table 2.2: Colony PCR conditions for amplification of *TbbCATL* in expression vectors

Cycle step	Temperature (ºC)	Duration (min)
Initial denaturation	95	5
Denaturation	95	1
Annealing	58	1
Extension	72	2

2.2.6 Plasmid isolation and restriction digestion

From the overnight culture, 1 ml was used to make glycerol stocks, stored at -80ºC and the remainder was used for plasmid isolation which was carried out using the GeneJet plasmid mini-prep kit according to the manufacturer's instructions. The purified plasmid was double digested with BamHI and XhoI in a 30 µl reaction volume at 37ºC for 1 hour. The enzymes were inactivated by incubation at 70ºC for 20 min. A 10 µl sample from the reaction was electrophoresed on a 1% (w/v) agarose gel and the remaining sample was electrophoresed on a 1% (w/v) agarose gel containing 10 µg/ml of crystal violet. The gene insert was extracted from the latter gel using the Zymo Research DNA gel recovery kit, as per the manufacturer's instructions.

2.2.7. Ligation of the *TbbCATL* gene into pET32a, pET28a and pGEX-4T expression vectors and transformation into *E. coli* BL21 (DE3) cells

A sample of the genes were sent for sequencing at the University of Stellenbosch Central DNA sequencing Facility and the insert from PTZ57R/T was chosen for further use as it has fewer mutations. The ligation was carried out in a 5:1 ratio of gene inserts to expression vector (Fig. 2.5) using T4 DNA ligase in a 30 µl reaction, at 37ºC for 1 hour followed by RT for 16 hours. The ligation mixtures, *TbbCATL*_pET32a, *TbbCATL*_pET28a and *TbbCATL*_pGEX-4T were transformed into *E. coli* BL21 (DE3) cells. Briefly, competent *E. coli* BL21 (DE3) cells were thawed on ice, 2 µl of the ligation mixture was added to 80 µl of the competent cells and incubated on ice for 30 min. The cells were heat shocked at 42ºC for 45 s and returned to ice for 2 min. The cells were added to 1 ml of pre-warmed 2xYT medium without antibiotic and incubated for 1 hour at 37ºC with shaking (200 rpm). The cells were spread onto pre-warmed 2xYT agar containing ampicillin for pGEXT-4T and pET32a, or kanamycin for pET28a and

incubated at 37°C overnight. Colony PCR was carried out on randomly selected colonies using the conditions described in Table 2.2. Glycerol stocks were made and stored at -80°C.
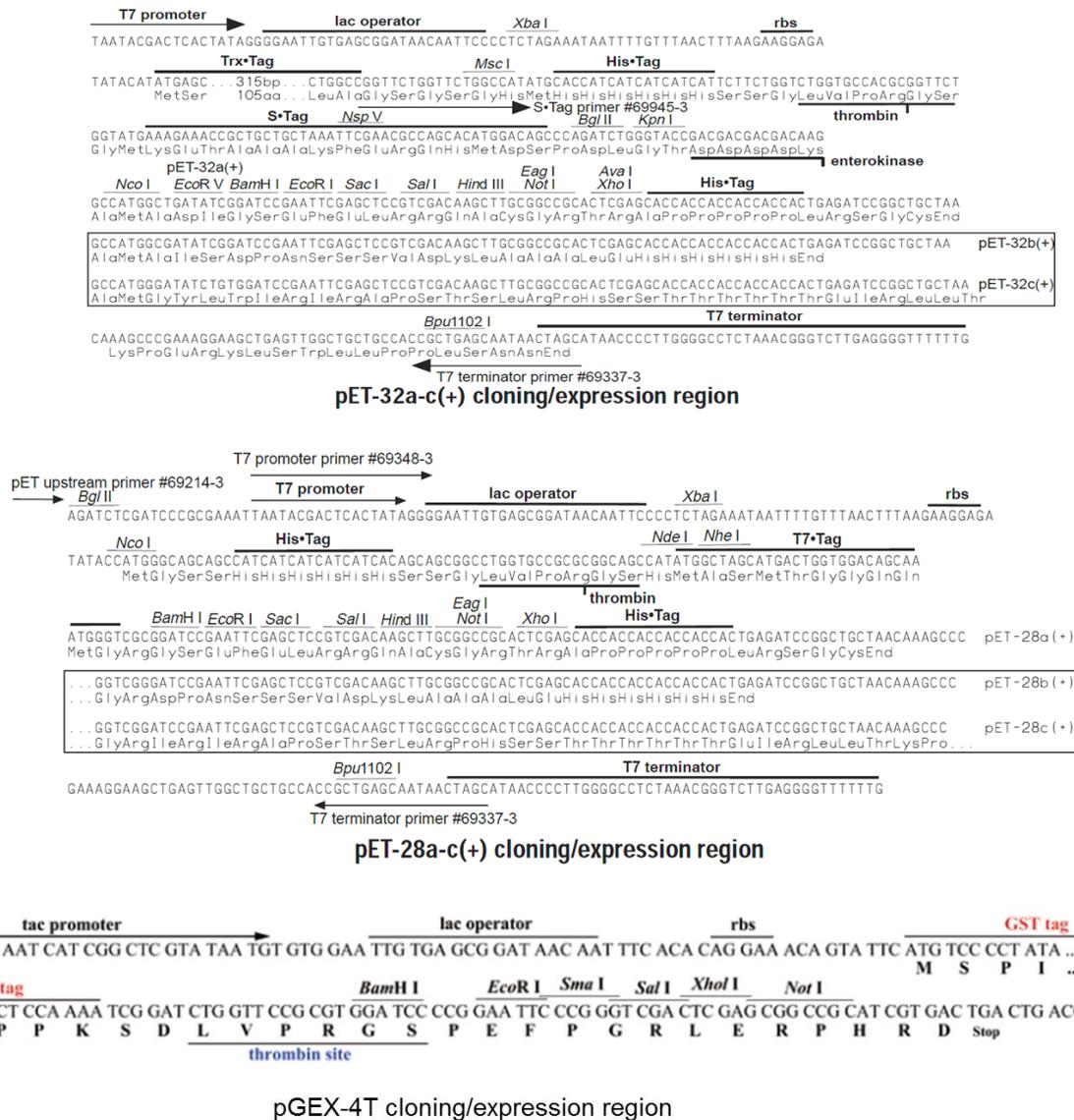
Figure 2.5: Cloning regions of pET32a, pET28a and pGEX-4T vectors. (Novagen and Genscript technical manuals).

2.2.8 Recombinant expression of *TbbCATL*_pET32a, *TbbCATL*_pET28a and *TbbCATL*_pGEX-4T in *E. coli* BL21 (DE3) cells

Glycerol stocks made from the colonies containing recombinant plasmids after sub-cloning of *TbbCATL* in *E. coli* BL21 (DE3) were streaked onto 2xYT medium containing 1% (w/v) bacteriological agar and 50 µg/ml ampicillin and incubated at 37°C overnight. A single colony was inoculated into 5 ml of 2xYT broth medium containing 50 µg/ml of ampicillin and incubated

at 37ºC with shaking at 200 rpm for 16 hours. A 1:100 dilution of the culture was done using fresh 2xYT medium (4 x 50 ml) and grown until it reached an OD_{600} of 0.6. Recombinant expression was induced with 100 mM IPTG for 4 hours, followed by centrifugation (6000 x g, 4ºC, 10 min). The pellets were combined and resuspended in 8 ml of PBS containing 1% (v/v) Triton-X-100 and 1 µg/ml lysozyme and incubated at 37ºC for 1 hour. The pellet was frozen at -70ºC and thawed at RT. The lysed pellet was sonicated five times for 20 s on ice and centrifuged (10 000 x g, 4ºC, 10 min). The rest of the bacterial lysate was kept for purification while a 50 µl sample was analysed by SDS-PAGE. The expected size of *Tbb*CATL including the purification tag in pET28a was 38 kDa, pET32a was 50 kDa and pGEX-4T was 58 kDa (Fig. 2.6).

Figure 2.6: Schematic representation of the expected *Tbb*CATL sizes in the expression vectors. Recombinant expression in pET28a includes a His tag; pET32a includes a thioredoxin (Trx) and a His tag; pGEX-4T includes a GST tag. The S-tag is depicted by the yellow box

2.2.9. *Tbb*CATL recombinant expression analysis by 12% reducing SDS-PAGE

The expression samples were analysed on SDS-PAGE according to the method described by (Laemmli, 1970). A sample (50 µl) from the bacterial lysate was combined with an equal volume of reducing treatment buffer [125 mM Tris-HCl, 4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) 2-mercaptoethanol] and boiled for 5 min followed by centrifugation (10 000 x g, 5 min, RT) using a benchtop centrifuge. The pellet was resuspended with an equal volume of PBS and reducing treatment buffer. A total of 10 µl of each sample was loaded onto a duplicate 12% reducing SDS-PAGE gels (Laemmli, 1970) and electrophoresed at 100V, 20A per gel in tank buffer [250 mM Tris-HCl buffer, pH 8.3, 192 mM glycine, 0.1% SDS] using the BioRad® Mini protein electrophoresis apparatus (BioRad, USA). One gel was stained with Coomassie blue R250 and the other gel used for a western blot. The images were captured using the G:Box system (Syngene, Cambridge UK). A standard curve for the molecular weight marker separated on a 12% SDS-PAGE was constructed (Fig. 2.7).
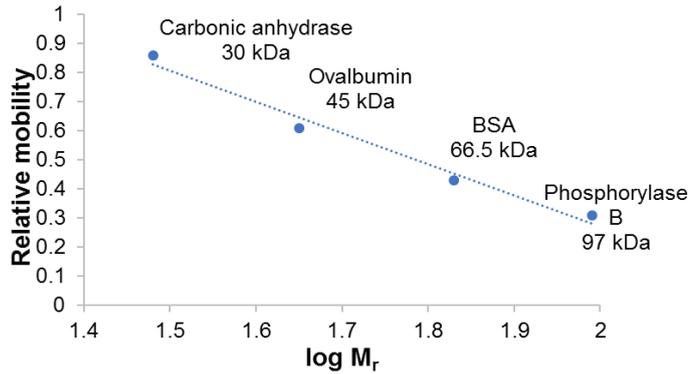
Figure: 2.7 Standard curve for relative mobility against molecular weight marker used in 12% SDS-PAGE. The molecular weight marker was made up of phosphorylase B (97.4 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa) and carbonic anhydrase (30 kDa). The trendline equation was given by y=-1.0718x+2.4147 with a correlation coefficient of 0.98.

2.2.10 Western blot analysis of *Tbb*CATL recombinant expression.

After reducing SDS-PAGE, the gel was equilibrated with blotting buffer [45 mM Tris-HCl buffer, pH 8.3, 173 mM glycine and 0.1% (w/v) SDS] for 15 min with gentle shaking at RT. The gel was transferred onto a nitrocellulose membrane using blotting buffer for 16 hours, 20A. The unoccupied sites on the nitrocellulose were blocked with 5% (w/v) low fat milk in Tris buffered saline (TBS; 20 mM Tris-HCl buffer, pH 7.4, 200 mM NaCl) for 1 hour at RT with gentle shaking followed by a washing step with TBS. The nitrocellulose membrane was incubated with chicken anti-GST IgY (1 µg/ml) for 4 hours at RT, followed by a washing step with TBS before incubation with rabbit anti-IgY-HPRO conjugate for 1 hour at RT. Finally, the membrane was incubated in substrate solution [0.06% (w/v) 4-chloro-1-naphthol, 0.1% (w/v) methanol, 0.0015% (v/v) H_2O_2 in TBS] and incubated in the dark for 3 min and the image captured using the G:Box system (Syngene, Cambridge UK).

2.2.11 Purification of recombinant *Tbb*CATL

Glutathione-agarose resin was suspended in distilled water and allowed to swell to a volume of 1 ml (max) at RT for 1 hour. The resin was transferred into a 10 ml BioRad® column and washed with 20 ml PBS followed by equilibration with 20 ml PBS containing 1% (v/v) Triton X-100. The *E. coli* lysate was transferred to the column and incubated on an end-over-end rotor at 4ºC overnight. The column was washed with 20 ml of PBS-Triton-X, and 1 ml fractions were collected. The protein was eluted with reduced glutathione buffer (50 mM Tris-HCl, 10 mM reduced glutathione, pH 9.5). The column was regenerated with 5 ml sodium borate buffer (200 mM borate in 500 mM NaCl pH 8.0) and washed with 5 ml dH₂O, followed by 5 ml sodium acetate buffer (100 mM acetate, 500 mM NaCl, pH 4.0) and 5 ml dH₂O and stored in 40% (v/v)

ethanol at 4ºC. The eluted fractions were analysed on a 12% reducing SDS-PAGE gel and stained with Coomassie blue R250 (Section 2.2.9).

2.2.12 Recombinant expression and purification of cathepsin L-like cysteine proteases from *T. congolense* (*Tco*CATL) and *T. vivax* (*Tvi*CATL).

*Tco*CATL and *Tvi*CATL genes were previously cloned into pPIC9 yeast expression vectors and transformed into *P. pastoris* GS115 (Pillay *et al.*, 2010; Eyssen *et al.*, 2018). The glycerol stocks were streaked onto yeast extract peptone dextrose (YPD) plates [1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose, 1% (w/v) bacteriological agar containing 10 µg/ml tetracycline] and incubated at 30 °C for three days. A single colony was inoculated into 100 ml of YPD liquid medium [1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose containing 10 µg/ml tetracycline] and incubated for 2 days at 30ºC with shaking (200 rpm). The culture was added to 450 ml of buffered medium glycerol yeast (BGMY) [1% (w/v) yeast extract, 2% (w/v) peptone, 100 mM potassium phosphate buffer pH 6.5, 1.3% (w/v) yeast nitrogen base (YNB) without amino acids, containing 10 µg/ml tetracycline and 50 µg/ml ampicillin] and incubated in baffled flasks at 30ºC for 3 days with shaking (200 rpm). The BGMY culture was centrifuged (2000 x *g*, 4ºC, 10 min) and the resulting pellet resuspended in 500 ml buffered minimal media (BMM) [100 mM potassium phosphate buffer pH 6.5, 1.3% (w/v) YNB, 0.0004% (w/v) biotin, 0.5% methanol, containing 10 µg/ml tetracycline and 50 µg/ml ampicillin]. The baffled flasks were covered with three layers of sterile cheese cloth, incubated as previously and supplemented with 0.5% methanol every 24 hours for seven days. The culture was centrifuged (2000 x *g,* 4ºC, 10 min) and the resulting supernatant was filtered through Whatman number 4 filter paper in preparation for protein purification.

The recombinant proteins were purified by three phase portioning (TPP) (Pike and Dennison, 1989). To the filtered supernatant, 30% (v/v) of tertiary butanol was added and mixed thoroughly. Ammonium sulfate was added to the mixture, 30% (w/v) for *Tco*CATL and 40% (w/v) for *Tvi*CATL and stirred until dissolved. The mixture was centrifuged (6000 x *g*, 4ºC, 10 min) in a swing out rotor. The interfacial supernatant was carefully removed and the remaining TPP mixture was repeatedly centrifuged to ensure complete separation from the other phases and the pellets were resuspended in PBS.

2.2.13 Purification of *Tco*CATL and *Tvi*CATL by molecular exclusion chromatography (MEC).

The proteins were further purified by molecular exclusion chromatography (MEC). The HiPrep® 16/16 Sephacryl® S-200 HR (16 x 600 mm) pre-packed column was prepared according to the manufacturer's instruction manual and connected to the ÄKTA prime plus

purifier. The column was equilibrated with 60 ml of Milli-Q® distilled water, followed by two column volumes (140 ml) of MEC buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0) at 1 ml/min flow rate and 0.2 MPa pressure. The protein sample was filtered through a 0.22 µm filter and injected into the column (0.5 ml/min, 0.2 MPa, 120 ml). An elution profile (Fig. 2.8) was previously done by Dr Rob Krause to construct a calibration curve to determine the relationship between the molecular weight and elution volume of proteins of known size and the newly purified proteins. Fractions (2 ml) were collected and analysed by 12% reducing SDS-PAGE gel and silver stained (Section 2.2.14).
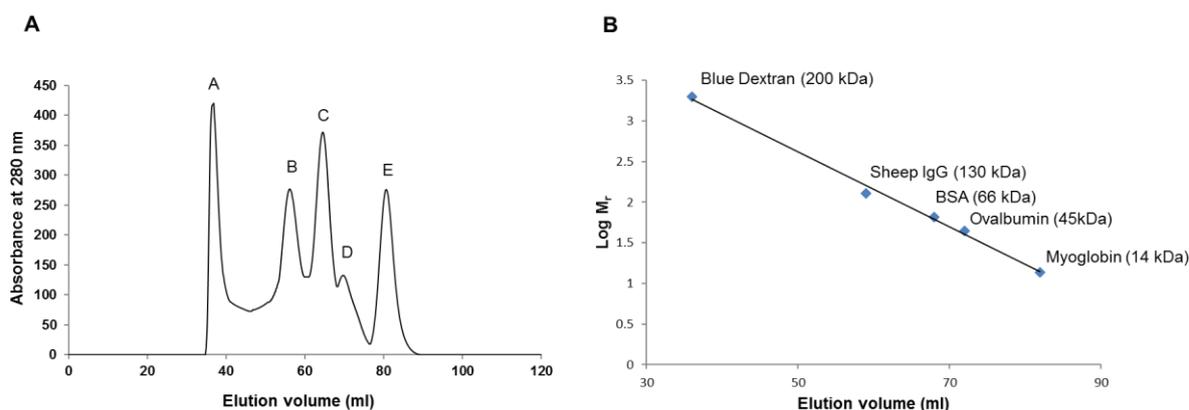
Figure 2.8: Elution profile for the calibration of the MEC column and Fischer's plot for protein size estimation. A, The standard protein mix was loaded onto the Sephacryl S-200 HR column (0.5 ml/min, 0.2 MPa, 120 ml) and eluted in MEC buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0), A, blue dextran (200 kDa), B: sheep IgG (130 kDa), C: BSA (66 kDa) and D: ovalbumin (45 kDa), E: myoglobin (14 kDa). **B**, the elution volumes of the standard proteins were plotted against their log molecular weight to construct the Fischer's plot calibration curve with the trend line equation of y=-0.0462x+4.9321 and correlation coefficient of 0.99.

2.2.14 Silver staining

Silver staining is a sensitive staining method used to visualise small amounts of proteins. This was carried out according to the procedure described by Nesterenko *et al.* (1994). Following SDS-PAGE, the gel was placed in a clean glass container and incubated in a fixation solution [0.8% (w/v) trichloroacetic acid (TCA) and 0.015% (v/v) formaldehyde in 50% (v/v) acetone] for 5 min at RT with gentle shaking. The gel was washed with dH_2O for 5 min before incubation with 50% acetone for 5 min at RT with gentle shaking followed by incubation in pre-treatment solution [0.01% (w/v) $Na_2S_2O_3 \cdot 5H_2O$] and washing with dH_2O. The gel was incubated in impregnate solution [0.26% (w/v) $AgNO_3$ and 0.37% (v/v) formaldehyde] for 8 min at RT with gentle shaking, washed with dH_2O for 10 s and incubated in developing solution [0.02% (w/v) Na_2CO_3, 0.015% (v/v) formaldehyde, 0.004% (w/v) $Na_2S_2O_3 \cdot 5H_2O$] until protein bands were

visible and immediately immersed in stopping solution [1% (v/v) glacial acetic acid] and rinsed with dH$_2$O.

2.2.15 Active site titration of GST-*Tbb*CATL, *Tco*CATL and *Tvi*CATL.

In order to determine the active concentration of the purified GST-*Tbb*CATL, *Tco*CATL and *Tvi*CATL, active site titration was conducted according to Barrett and Kirschke (1981). The proteases [1 µM diluted in 0.1% (w/v) Brij-35] were incubated with different concentrations of the cysteine protease inhibitor E-64 (0-1 µM) diluted in 0.1% (w/v) Brij-35 in assay buffer [100 mM Bis-Tris, pH 6.5, 4 mM EDTA, 0.02% (w/v) NaN$_3$, 6 mM DTT] and incubated for 30 min at 37°C. The substrate Z-Phe-Arg-AMC (20 µM) for GST-*Tbb*CATL and *Tco*CATL or H-D-Val-Leu-Lys-AMC (20 µM) for GST-*Tbb*CATL and *Tvi*CATL was added and fluorescence (Ex$_{360nm}$; Em$_{460nm}$) was measured using the FLUOStar Optima spectrophotometer (BMG Labtech, Germany).

2.3 Results

2.3.1 Cloning of the *TbbCATL* gene into the pGEM®-T and pTZ57R/T cloning vectors

The 954 bp gene coding for the catalytic domain of the *T. b. brucei* cathepsin L-like cysteine protease (*Tbb*CATL) was identified from the *T. b. brucei* genome. The gene was amplified by PCR using genomic DNA and showed a band of *ca*. 1000 bp (Fig. 2.9) which corresponds to the size of the target gene.
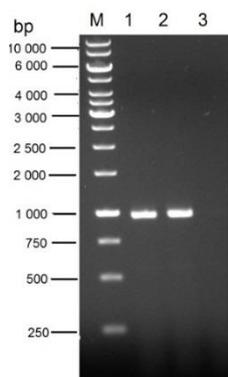
Figure 2.9: Agarose gel analysis of the *TbbCATL* PCR product amplified from *T. b. brucei* genomic DNA. *TbbCATL* gene was amplified with gene specific primer and analysed on a 1% agarose gel. M, molecular weight marker; lanes 1 and 2, *TbbCATL* PCR product; lane 3, non-template control.

The gene fragment was extracted from the agarose gel and ligated into both pGEM®-T and PTZ57R/T cloning vectors and subsequently transformed into *E. coli* JM109 cells. To confirm the presence of the *Tbb*CATL gene fragment in the vectors, the gene was amplified by PCR using specific gene primers in pGEM®-T and PTZ57R/T vectors. Analysis of the PCR product

on an agarose gel showed a band at *ca.* 1000 bp (Fig. 2.10, panel A) which correspond to the target gene's size. Following isolation of the plasmids from the *E. coli* cells and restriction enzyme digestion with BamHI and XhoI restriction enzymes, the gene bands were observed at *ca.* 1000 bp, and plasmid bands were observed at *ca.* 3000 bp for both pTZ57R/T and pGEM-T as expected (Fig. 2.10, panel B).
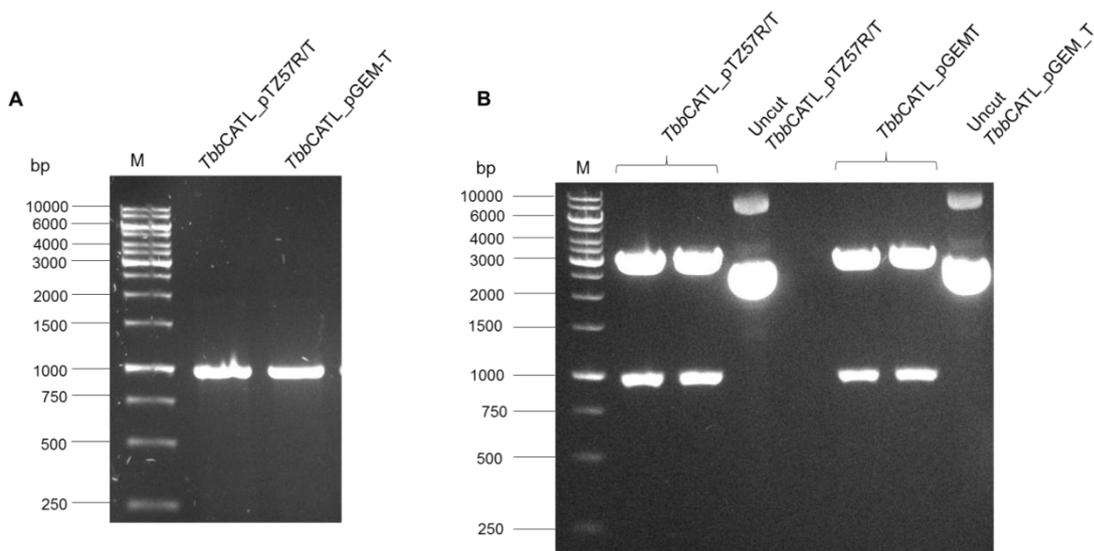
Figure 2.10: Colony PCR of *TbbCATL*_PTZ57R/T and *TbbCATL*_pGEM®-T in *E. coli* JM109 and restriction digestion analysis on 1% agarose. A, the *TbbCATL* gene insert was amplified specific from *TbbCATL*_PTZ57R/T and *TbbCATL*_pGEM®-T using gene specific primers. **B**, Recombinant *TbbCATL*_PTZ57R/T and *TbbCATL*_pGEM®-T plasmids were digested with BamHI and XhoI and compared to undigested plasmids. M, molecular weight marker.

2.3.2 Transformation of recombinant *TbbCATL*_pET28a, *TbbCATL*_pET32a and *TbbCATL*_pGEX-4T expression vectors into *E. coli* JM109 cells.

The recombinant plasmids were transformed into *E. coli* JM109 cells. The success of transformation was determined by analysis of the amplification product using specific gene primers. *Tbb*CATL_pET28 and *Tbb*CATL_pGEX-4T showed amplification *ca.* 1000 bp (Fig. 2.11, panels A and B respectively) which corresponds to the target gene size. No amplification was observed from *Tbb*CATL_pET32a (result not shown).
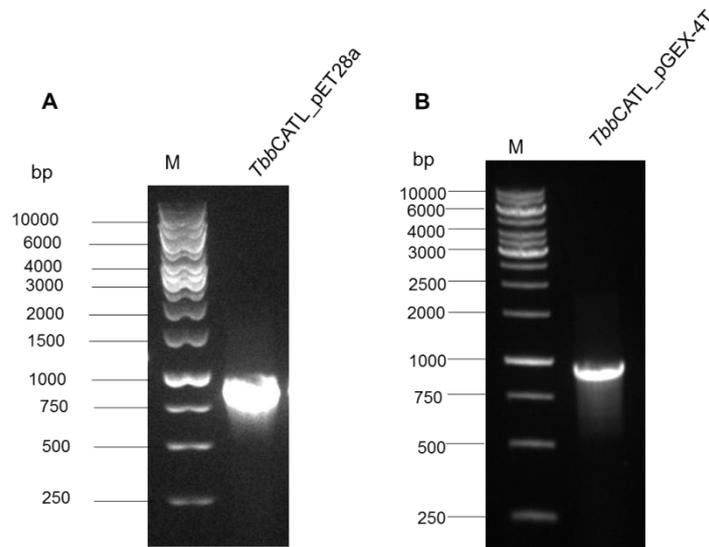
Figure 2.11: Analysis of *TbbCATL* gene amplification in recombinant pET28a and pGEX-4T after transformation into *E. coli* JM109 cells. Randomly selected colonies were amplified with gene specific primers in colony PCR and analysed on a 1% agarose gel**. A**, *TbbCATL*_pET28a *E. coli* transformant. **B**, *TbbCATL*_pGEX-4T *E. coli* transformant. M; molecular weight marker

A sample of each plasmid was subjected to a small-scale restriction digestion using BamHI and XhoI as done previously and subsequently analysed on an agarose gel. For each vector a band was obtained at ca. 1000 bp and another band *ca* 3000 bp for pET32a, 6000 bp for pET28a and 5000 bp for pGEX-4T vectors (Fig. 2.12, panels A, B and C), which further confirmed the presence of the *TbbCATL* gene insert in the expression vectors.
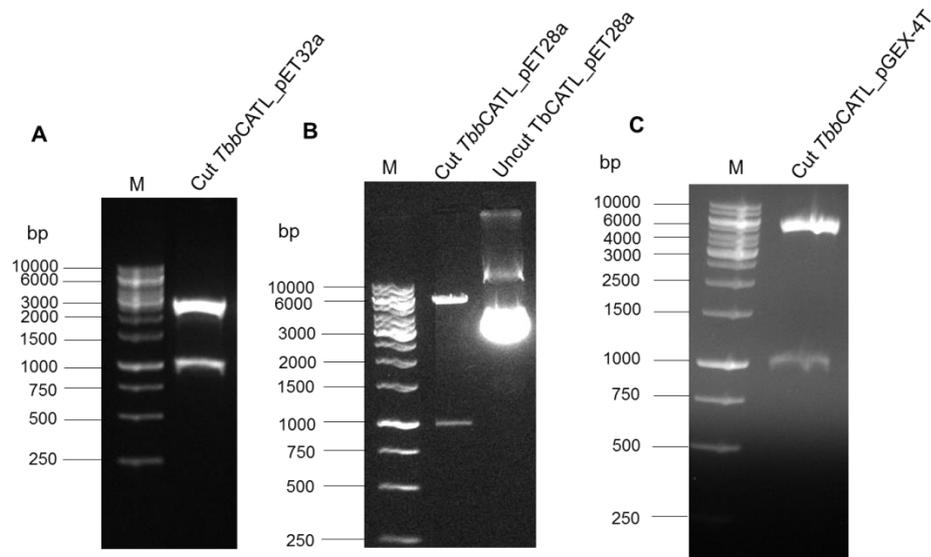
Figure 2.12: Analysis of restriction digestion of the *TbbCATL* gene from *TbbCATL*_pET32a, *TbbCATL*_pET28a and *TbbCATL*_pGEX-4T clones using BamHI (fwd) and XhoI (rev). The samples were electrophoresed on a 1% agarose gel. **A**, restriction digestion of *TbbCATL* gene insert from pET32a. **B**, restriction digestion of the *TbbCATL* gene insert from pET28a, **C**, Restriction digestion of the *TbbCATL* gene insert from pGEX-4T. M; molecular weight marker.

2.3.3 Transformation of *TbbCATL*_pET32a, *TbbCATL*_pET28a and *TbbCATL*_pGEX®_4T recombinant clones into E. coli BL21 (DE3) cells.

The plasmids isolated from *E. coli* JM109 cells were transformed into *E. coli* BL21 (DE3) cells. Eight colonies were randomly selected and used as templates in colony PCR using T7 promoter and terminator vector primers and specific gene primers and analysed by agarose gel electrophoresis. Amplification was only obtained with gene specific primers for *TbbCATL*_pET32a at *ca*. 1000 bp and *TbbCATL*_PGEX-4T between 500 and 750 bp (Fig. 2.13, panels A and C respectively). The *TbbCATL* gene in the pET28a vector was amplified with vector primers at *ca*. 1000 bp and gene specific primers between 500 and 750 bp (Fig. 2.13, panel B).
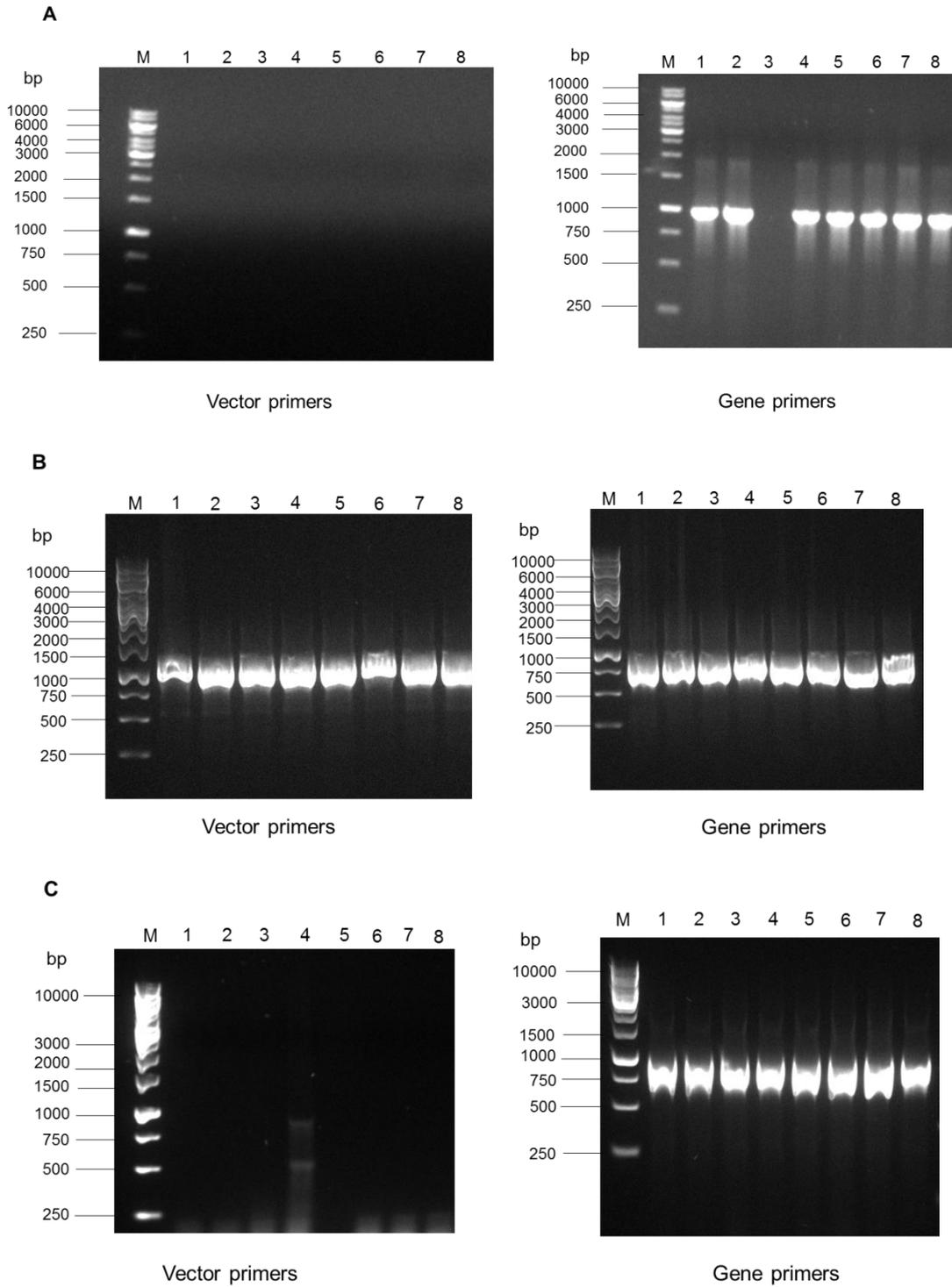
Figure 2.13: Agarose gel analysis of *E. coli* BL21 (DE3) transformed with *TbbCATL*_pET32a, *TbbCATL*_pET28a and *TbbCATL*_pGEX®_4T recombinant clones. Eight colonies were randomly selected and amplified with either T7 promoter and terminator vector primers (left hand panels) or gene specific primers (right hand panels) in a colony PCR. **A**, *TbbCATL* in pET32a; **B,** *TbbCATL* in pET28a; **C,** *TbbCATL* in pGEXT-4T. M, molecular weight marker.

45

2.3.4 Recombinant expression of *Tbb*CATL in *E. coli* BL21 (DE3) cells.

Analysis of recombinant expression of *Tbb*CATL in *E. coli* BL21 (DE3) cells by reducing SDS-PAGE (Fig. 2.14) showed that there were no prominent protein bands observed for *Tbb*CATL_pET32a (expected at 50 kDa) (Fig. 2.14, panel A). A prominent band was observed for *Tbb*CATL_pET28a expression at 66.5 kDa (expected at 38 kDa) (Fig. 2.14, panel B), but this band was not recognised by anti-His-tag antibodies (see below). The samples were also passed through a nickel column and no protein binding was observed. A prominent band was also observed at *ca.* 60 kDa when *Tbb*CATL was expressed in pGEX-4T (Fig. 2.14, panel C). Analysis of the soluble and insoluble fractions showed that GST-*Tbb*CATL was expressed in the soluble fraction (Fig. 2.14, panel C).

A western blot was done to detect the His tag in *Tbb*CATL_pET32a and *Tbb*CATL_pET28a, and the GST tag from *Tbb*CATL_pGEX-4T. There were no bands detected by the mouse anti-6xHis-HRPO antibody (Fig. 2.14, panels A and B). The GST tag was detected in all the pGEX-4T expression fractions, including the non-induced fraction of *Tbb*CATL_pGEXT-4T (Fig. 2.14, panel C). A prominent band at *ca.* 60 kDa was observed in the induced and soluble fractions, which confirms soluble expression of GST-*Tbb*CATL in pGEX-4T.
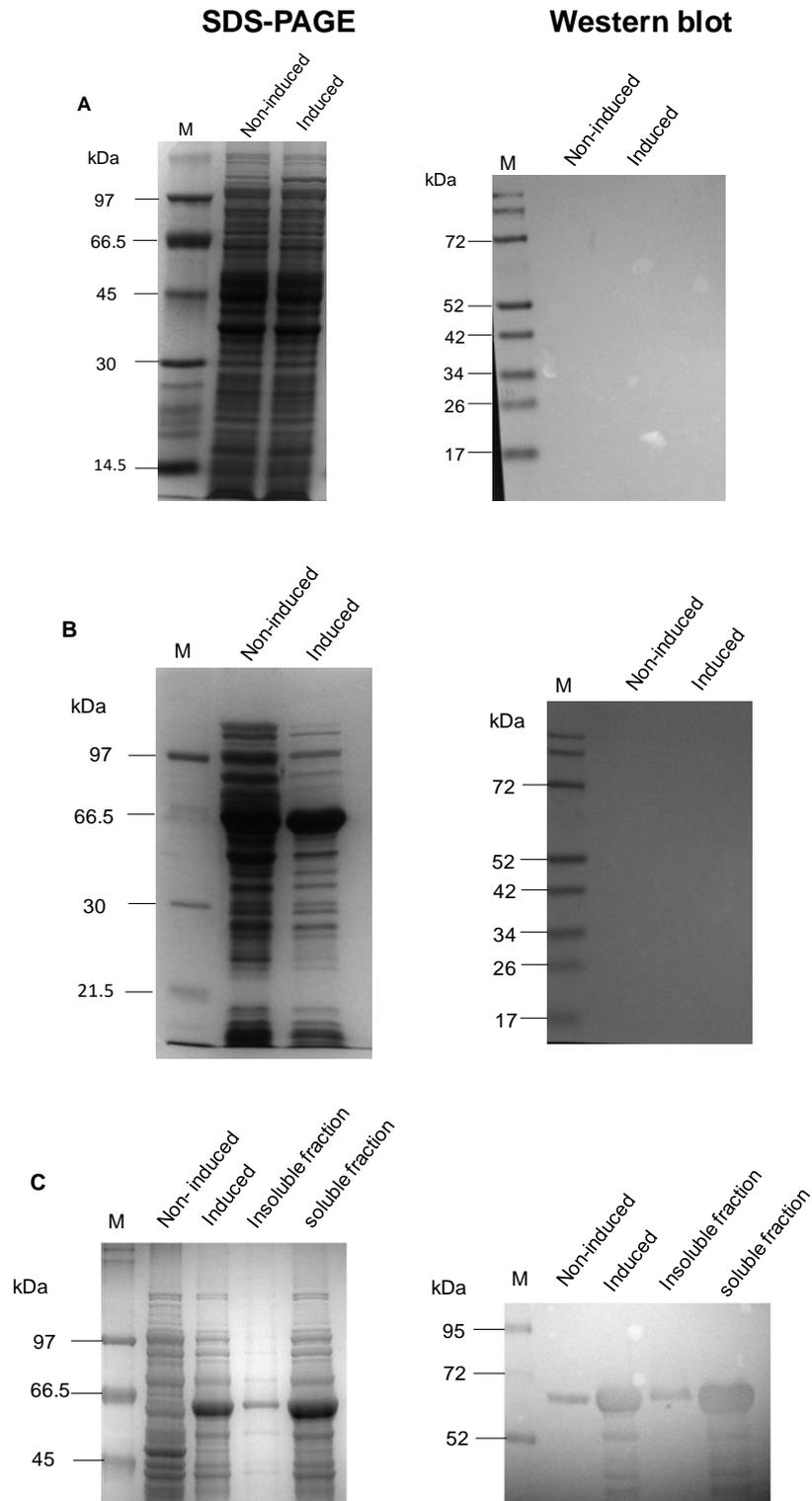
Figure 2.14: Analysis of the recombinant expression of *Tbb*CATL on a 12% reducing SDS-PAGE gel and by western blot. Recombinant expression was induced in **A**, pET32a; **B**, pET28a and **C**, pGEX-4T. Samples were analysed on a 12% reducing SDS-PAGE, stained with Coomassie blue R-250. In a western blot, proteins were transferred onto a nitrocellulose and incubated with mouse anti-6xHis HRPO conjugate (Panels A and B); anti-GST IgY (Panel C) (primary antibody) followed by rabbit anti-IgY- HPRO (secondary antibody) and H_2O_2/4-chloro-1-naphthol. M, molecular weight marker.

2.3.4.1 Purification of GST-*Tbb*CATL using glutathione agarose affinity chromatography

Following recombinant expression of soluble GST-tagged *Tbb*CATL (Fig. 2.14, panel C), the crude bacterial lysate was purified using glutathione agarose affinity chromatography and the eluted fractions were analysed by reducing SDS-PAGE (Fig. 2.15). Recombinant GST-*Tbb*CATL was eluted as a single band of approximately 60 kDa in fractions 1 to 10 (Fig 2.15, panel B) with reduced glutathione buffer.
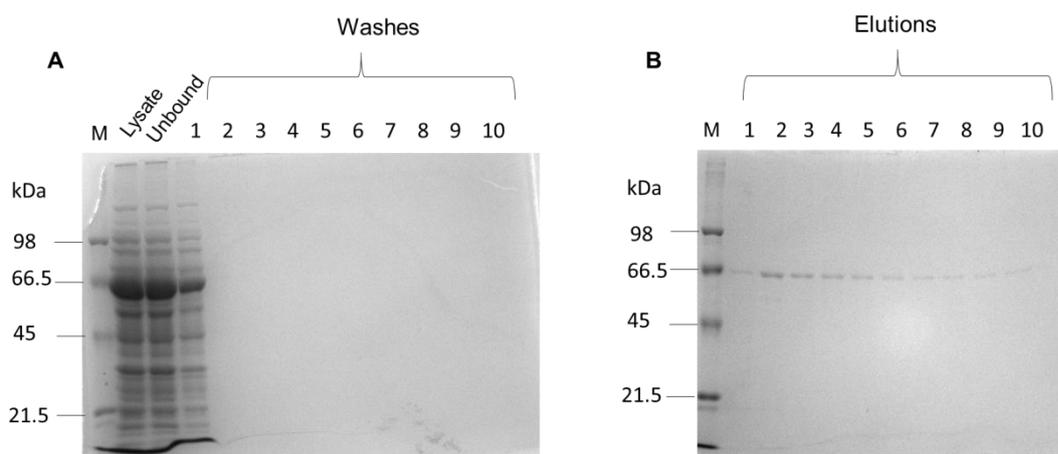
Figure 2.15: Analysis of purification of GST-*Tbb*CATL by GST affinity chromatography analysed on a 12% reducing SDS-PAGE gel. The samples were collected in 1 ml fractions at RT. **A**, PBS-Triton X-100 was used for binding of protein to the column and the washing step. **B**, Elution of protein with reduced glutathione buffer [50 Mm Tris-HCl buffer, pH 9, containing 10 mM reduced glutathione]. M, molecular weight marker.

2.3.5 Recombinant expression and purification of the catalytic domain of *T. congolense* (*Tco*CATL) and *T. vivax* (*Tvi*CATL)

*Tco*CATL and *Tvi*CATL were recombinantly expressed in *P. pastoris* yeast cells over two rounds of expression and recombinantly expressed proteins were analysed by reducing SDS-PAGE (Fig. 2.16). *Tco*CATL was observed at the expected size of approximately 29 kDa. After round two of expression, the intensity of this band increased, indicating increased levels of expression. Two additional bands were observed between 31 kDa and 66.5 kDa. *Tvi*CATL was observed as two bands of approximately 28 kDa and 32 kDa.
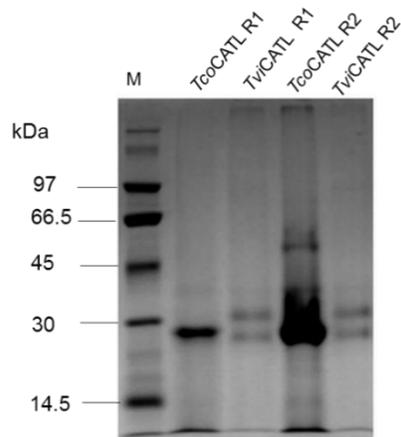
Figure 2.16: Recombinant expression of *Tco*CATL and *Tvi*CATL analysed by 12% reducing SDS-PAGE. The proteins were purified by three phase partioning prior to analysis. R1, round 1 of expression; R2, round 2 of expression; M, molecular weight marker.

2.3.5.1 Purification of recombinant *Tco*CATL and *Tvi*CATL by molecular exclusion chromatography (MEC)

The recombinant *Tco*CATL and *Tvi*CATL were further purified by Sephacryl S-200 MEC. For *Tco*CATL, protein absorbance peaks were observed at 40 ml and between 62 and 90 ml (Fig 2.17, panel A) which corresponds to the expected size of 29 kDa. For *Tvi*CATL, protein absorbance peaks were observed at 40 ml and between 60 ml and 100 ml (Fig. 2.18, panel A) which correspond to the expected sizes of 28 kDa and 32kDa. Analysis by reducing SDS-PAGE and silver staining showed a single band corresponding to the size of *Tco*CATL (30 kDa) in fractions 72 to 86 ml (Fig. 2.17, panel B). The two bands at 28 kDa and 32 kDa corresponding to *Tvi*CATL were eluted together in fractions 80 to 88 ml (Fig. 2.18, panel B).
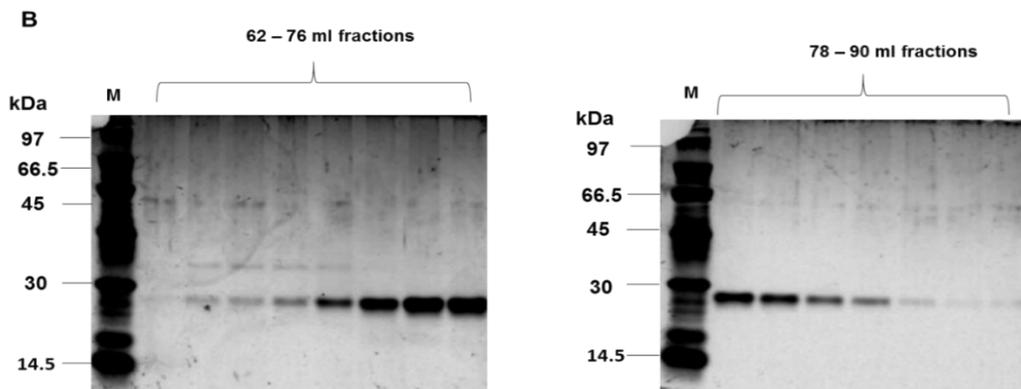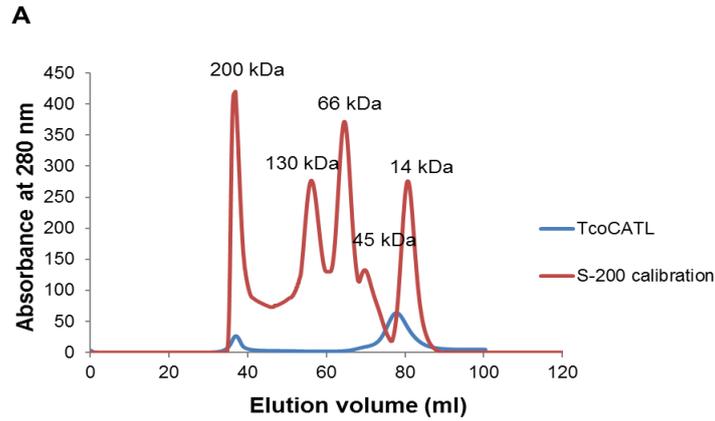
Figure 2.17: Purification of *Tco*CATL on a Sephacryl® S-200 HR molecular exclusion chromatography column. *Tco*CATL was loaded onto the column (16 x 600 mm, flow rate of 0.5 ml/min, 2 ml/fraction, 120 ml) and eluted with MEC buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, pH 8.0). **A,** Elution profile of the 2 ml fractions at 280 nm. **B**, analysis of fractions 62 – 90 ml a 12% reducing SDS-PAGE gel followed by silver staining. M; molecular weight marker.
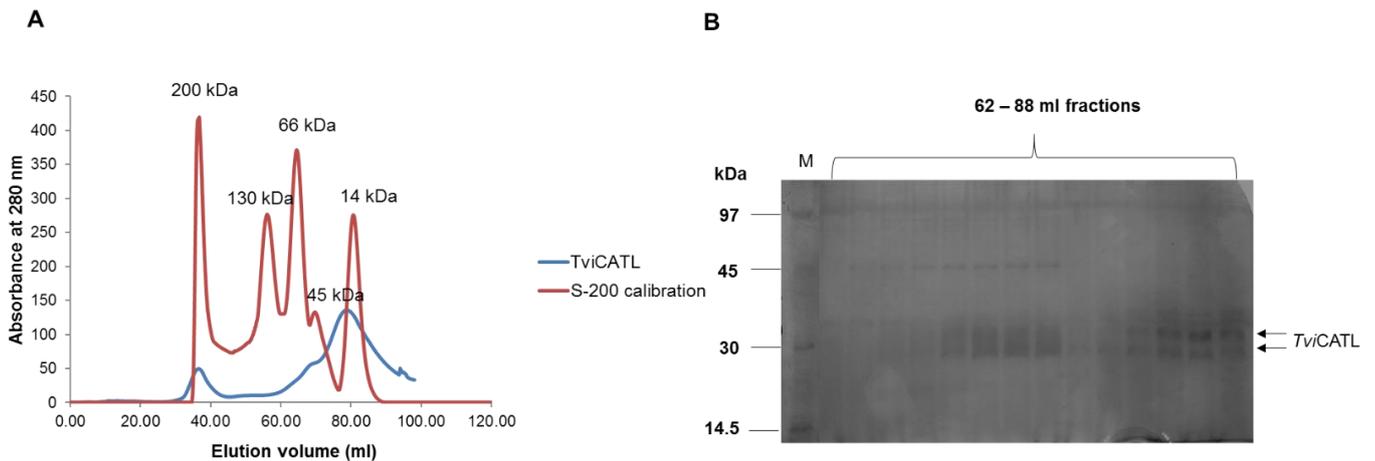
Figure 2.18: Purification of *Tvi*CATL on a Sephacryl® S-200 HR molecular exclusion chromatography column. *Tvi*CATL was loaded onto the cloumn (16 x 600 mm, flow rate of 0.5 ml/min, 2 ml/fractions, 120 ml) and eluted with MEC buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, pH 8.0). **A,** Elution profile of the 2 ml fractions at 280 nm. **B**, analysis of fractions 62 – 88 ml on a 12% reducing SDS-PAGE gel followed by silver staining. M; molecular weight marker.

To estimate the active concentrations of the purified samples, an active site titration was done with E-64. Residual activity of GST-*Tbb*CATL and *Tco*CATL hydrolysis of Z-Phe-Arg-AMC and GST-*Tbb*CATL and *Tvi*CATL hydrolysis of H-D-Val-Leu-Lys-AMC substrates in the presence of E-64 was determined by measuring the absorbance of the released AMC. *Tco*CATL and *Tvi*CATL were able to cleave their respective substrates in the absence of E-64, while the addition of E-64 inhibited their activity in a dose dependant way (Fig. 2.19). Approximately 20% of *Tco*CATL was active which is equivalent to 0.2 µM (Fig. 2.19, panel A) and approximately 30% *Tvi*CATL was active which is equivalent to 0.3 µM (Fig. 2.19, panel B). No activity was observed for GST-*Tbb*CATL.
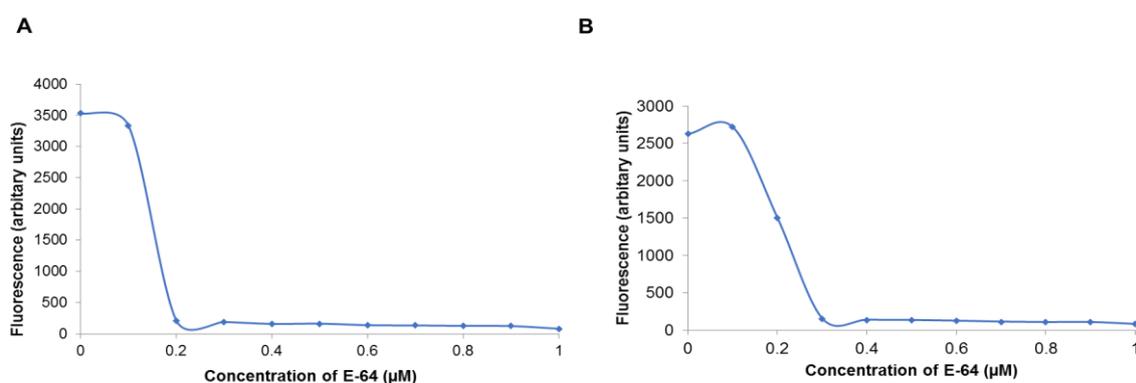
Figure 2.19: Active site titration of *Tco*CATL and *Tvi*CATL. *Tco*CATL and *Tvi*CATL (1 µM) were titrated against E-64 ranging from 0-1 µM and incubated for 30 min at 37°C before incubation with **A**, Z-Phe-Arg-AMC and **B**, H-D-Val-Leu-Lys-AMC respectively. The mean residual activity against the substrates was measured (each data point represents the mean activity of triplicate samples).

2.4 Discussion

Cysteine proteases of trypanosomes have been discovered to play a crucial role in the pathogenicity of the parasite (Sajid and McKerrow, 2002). They are immunodominant and predominantly expressed during the mammalian infective stage of their life cycle, therefore are attractive diagnostic and drug targets (Serveau *et al.*, 2003). In this study, cathepsin L-like cysteine proteases of the animal infective trypanosoma species *T. b.i brucei* (*Tbb*CATL), *T. congolense (Tco*CATL) and *T. vivax (Tvi*CATL*)* were recombinantly expressed and purified. Moreover, cloning and recombinant expression of *Tbb*CATL is reported for the first time in this study.

The *TbbCATL* gene sequence was identified and amplified from the Lister 427 strain of *T. b. brucei*, and ligated into pGEMT and PTZ57/R TA cloning vectors. The gene sequence in PTZ57R had fewer mutations hence the fragment was used in further sub-cloning procedures. The pET vectors with a 6xhistidine tag (His-tag) are usually the first choice for recombinant protein expression in *E. coli*. Despite the possibility of decreased solubility of the fusion protein,

the pET vectors are desirable because of high protein expression levels in *E. coli* hosts due to a strong promoter, the His tag is small (in pET28a) and requires a simple purification method such as the nickel affinity chromatography and does not interfere with downstream applications of the expressed protein (Rosano and Ceccarelli, 2014). In light of this, pET28a was the first vector system to be investigated for recombinant expression of *Tbb*CATL, followed by pET32a which has a Trx tag and a His-tag and lastly pGEX-4T which has a GST tag. Amplification with gene specific primers in a PCR showed that the DNA insert was only present in pET28a and pGEX-4T. A study by Álvarez *et al.* (2002) revealed that the use of gene specific primers may give false positives due to the direct amplification of non-cloned DNA fragments which enter the competent cells, together with recircularised plasmids, during transformation, hence, the presence of the *TbbCATL* gene in the vectors was confirmed with restriction digestion which showed that the gene insert was also present in pET32a.

There was no expression obtained with the pET vectors in this study even after optimisation of temperature, induction time and IPTG concentration. In previous studies, attempts to recombinantly express *Tcr*CATL using common vectors that use T7 and tac promoters was also unsuccessful, therefore this protease was expressed in the pCheYl5LOX expression plasmid. This vector expresses insoluble CheY fusion proteins which was required for the expression of significant amounts in the case of *Tcr*CATL (Eakin *et al.*, 1993). There are no reports on the use of pET vectors for the expression of *Tco*CATL.

Due to the failed expression in the pET vectors the pGEX-4T vector was used and high levels of *Tbb*CATL was expressed in soluble form. Expression was also observed in the non-induced sample which may be due to leaky expression. According to Terpe (2003), the majority of GST fusion proteins are expressed in the soluble form while others are partly soluble. The purification of the expressed GST-*Tbb*CATL fusion protein required optimisation of the lysis method and incubation time of the bacterial lysate with the glutathione-resin. However, subsequent attempts to express *Tbb*CATL resulted in an insoluble protein. Insoluble expression of GST fusion proteins is often associated with hydrophobic regions, many charged amino acid residues and large fusion proteins (≥100 kDa) (Terpe, 2003). In this case, however, factors that contributed to the insolubility of the protein remain unclear. Furthermore, attempts to purify the protease after solubilisation were unsuccessful, therefore subsequent studies, i.e. antibody production and ELISAs, were carried out using the GST tagged *Tbb*CATL.

Previous attempts to express *Tco*CATL in bacterial systems resulted in an incorrectly folded and inactive protease which led to the investigation and optimisation of expression in *P. pastoris* (Boulangé *et al.*, 2001). One of the main advantages of using the methylotrophic yeast *P. pastoris* is that recombinantly expressed proteins are secreted into the medium. The

secretion signal α-MF contains a pre-region which directs the protein to the endoplasmic reticulum (ER) and a pro-region which plays a role in the transfer of the protein from the ER to the Golgi apparatus. Secretion is efficient for the production and purification of proteins (Ahmad *et al.*, 2014).

In a study by Mendoza-Palomares *et al.* (2008), cathepsin B-like cysteine proteases, *Tco*CB, was recombinantly expressed in methyloptrophic yeast *P. pastoris* as a highly glycosylated enzyme and the non-glycosylated mutant forms were obtained by disruption of glycosylation sites. The full length and truncated form of *T. b. rhodesiense* cathepsin L-like cysteine proteases has also been expressed in *P. pastoris* (Caffrey *et al.*, 2001). In both studies, the pro-domain was included as it was shown to help with the proper folding of the enzyme. In the present study, the *P. pastoris* expression system was not investigated for *TbbCAT*L however may be an ideal expression system for this protease. The primers that would be required for this are: <u>forward primer</u> 5'- GC <u>CTCGAG</u> **ATG** GCG TGC CTT GCG TCT GTC G -3', <u>reverse primer</u> 5'- GC <u>GAATTC</u> **TCA** ACT GCG GAG ACG GCC TG -3' with restriction sites for XhoI (forward) and EcoR1 (reverse) which are underlined.

Trypanosomal cysteine proteases *Tco*CATL and *Tvi*CATL were expressed using the methylotrophic *P. pastoris* expression system. The expression was carried over two rounds and interestingly, the second round of expression yielded higher levels of *Tco*CATL. Generally, in this expression system, the growth phase is separated from the production phase, therefore the accumulated biomass is not stressed by the expressed protein. Nevertheless, during the production phase, the biomass continues to accumulate which may account for the higher levels of expressions of *Tco*CATL in the second round (Ahmad *et al.*, 2014). However, this also resulted in higher levels of expression of yeast proteins as depicted by additional bands on reducing SDS-PAGE gels.

*Tvi*CATL was expressed as two proteins bands (28 kDa and 32 kDa) which is consistent with the observation reported by Eyssen *et al.* (2018) where it is shown that the 32 kDA form is glycosylated. This protease is predicted to undergo N-glycosylation at Asn[288] (*Tvi*CATL numbering) resulting in a mixture of glycosylated (32 kDa) and non-glycosylated (28 kDa) *Tvi*CATL. Glycosylation was also reported in *Tcr*CATL and the glycosylation site is also found in the homologue found in *Trypanosoma brucei rhodesiense* (*Tbr*CATL). In *Tco*CATL, the Asn is replaced by an Asp residue which explains why this protease is not glycosylated.

In this study, *Tco*CATL and *Tvi*CATL were first purified by three-phase partitioning (TPP) before molecular exclusion chromatography (MEC). Generally, the contaminants are partitioned in the t-butanol and aqueous phase while the biomolecules are purified from the interphase during TPP (Dennison and Lovrien, 1997). In a study by Caffrey *et al.* (2001),

purification of *Tbr*CATL was carried out on the *P. pastoris* culture supernatent using a lyophilisation step followed by a clearing step using an acidic buffer to remove enzyme proforms and contaminants and lastly a buffer exchange step. The recombinantly expressed pro-enzyme *Tco*CATB included a His-tag which allowed for purification by immobilised metal ion affinity chromatography (IMAC). The enzyme was activated in a low pH buffer and the catalytic domain purified by ion-exchange chromatography (Mendoza-Palomares *et al.*, 2008).

Three phase partitioning has been reported to enhance the activity of other proteins while it denatures proteins with the quaternary structure such as haemoglobin. Moreover, this technique has been used in the purification of proteases from microbial and plant sources (Nitsawang *et al.*, 2006). *Tco*CATL and *Tvi*CATL were recovered at 30% and 40% (w/v) ammonium sulfate respectively, which is within the range of optimum parameters reported for other proteolytic enzymes (Gagaoua and Hafid, 2015).

The activity of the purified *Tco*CATL and *Tvi*CATL proteases was studied using fluorogenic peptide substrates and a class-specific protease inhibitor, E-64. Consistent with the findings reported by Eyssen *et al.* (2018), *Tvi*CATL was able to hydrolyse H-D-Val-Leu-Lys-AMC. This protease prefers basic amino acid residues Lys and Arg in P_1 position and bulky hydrophobic residues (Phe or Leu) in the P_2 position. These results also show that the glycosylation in *Tvi*CATL does not affect enzyme activity. The activity of *Tco*CATL was studied using the standard substrate Z-Phe-Arg-AMC. Studies have shown that *Tco*CATL prefers substrates with Pro in P_2, which is a characteristic of trypanosomal cysteine proteases that distinguishes them from their mammalian homologues (Boulangé *et al.*, 2001). E-64 was able to inhibit the activity of *Tco*CATL and *Tvi*CATL as expected. In a study by Troeberg *et al.* (1996), the native *Tbb*CATL succesfully cleaved the synthetic substrate Z-Phe-Arg-NHMec. In the present study, no activity was observed with recombinant *Tbb*CATL on the synthetic substrate Z-Phe-Arg-AMC. It is recommended that the fusion tag be removed from recombinant proteins prior to an activity assay as they interfere with the activity and structure of the protein (Rosano and Ceccarelli, 2014).

In conclusion, *Tbb*CATL, *Tco*CATL and *Tvi*CATL were recombinantly expressed, purified and subjected to an enzyme assay. *Tbb*CATL showed no activity while *Tco*CATL and *Tvi*CATL were able to hydrolyse synthetic substrates and were inhibited by a class specific inhibitor. For future studies, *Tbb*CATL would be recombinantly expressed using the methylotrophic yeast system *P. pastoris* without a tag, as done with *Tco*CATL and *Tvi*CATL.

Chapter 3

Anti-*Tbb*CATL antibody production in chickens and biopanning of single chain variable fragments from the Nkuku® phage library against recombinant *Tvi*CATL

3.1 Introduction

To achieve effective control of animal African trypanosomiasis, accurate diagnosis is essential. Diagnosis cannot rely on clinical symptoms alone as these are varied and may easily be mistaken for other diseases. Consequently, parasitological diagnostic techniques are often a necessity. However, these techniques are cumbersome and diagnosis is not immediate thus alternative diagnostic methods are sought (Büscher *et al.*, 2017). Although research towards development of a reliable diagnostic test for AAT has been on-going for decades, not much progress has been made due to limited economical interest in veterinary diagnostics, inconsistencies in research priorities and financial resources.

Improvement has been made with the development of molecular and serological diagnostic techniques such as polymerase chain reaction (PCR) detecting trypanosomal antigens and immunosorbent assays (ELISA). The latter is an indirect antibody detection test that has been used to detect *T. congolense*, *T. vivax* and *T. brucei* in infected animals. These ELISAs are sensitive, but not specific since the antigen used is usually a crude parasite lysate. Specific detection is dependent on the purity of the antigen therefore this technique lacks standardisation (Ezeani *et al.*, 2008). In addition, antibodies are detectable even months after the disease is cured, therefore, these tests are not suitable follow ups. Card agglutination test for trypanosomes (CATT), an antibody based detected test, has been successfully used to detect *T. b. gambiense* infections but unsuccessful for *T. b. rhodesiense* infections (Büscher *et al.*, 2013). Consequently, there is still a need for reliable and specific detection tools. Recently, a quick-format diagnostic test was developed for field diagnostics. The test is based on the recombinant cathepsin-B like cysteine protease from *T. congolense* (*Tco*CATB) and the tandem repeat protein (GM6) from *T. vivax* (*Tv*GM6). Although this test showed high specificity, *T. vivax* infected animals gave a signal for both the *Tco*CATB and GM6 repeat. Additionally, the proteases were detected between two to four weeks of infection which is longer that the classical ELISA which detects them at one to two weeks after infection, therefore requires improvement (Boulangé *et al.*, 2017).

Antigen based ELISA tests such as sandwich ELISA are ideal for a diagnostic test as they

allow the detection of circulating parasite antigen, thereby diagnosing current infections rather than antibodies (Boulangé *et al.*, 2017). Studies showed that trypanosomal cathepsin L-like cysteine proteases *Tco*CATL and *Tvi*CATL are released into the infected host bloodstream upon parasite lysis and is involved in pathogenesis during trypanosomal infections. Additionally, trypanotolerant cattle breeds challenged with *Tco*CATL, were reported to produce a prominent anti-*Tco*CATL IgG response which was not present in a susceptible cattle breed (Authié *et al.*, 2001). Similarly, antibodies could be detected in the serum of cattle experimentally challenged with *T. vivax* and it was possible to discriminate infected and non-infected sera in an ELISA based on recombinant *Tvi*CATL (Eyssen *et al.*, 2018). Therefore, these proteases are promising targets in the development of diagnostic tools. Antigen based ELISAs require good antibodies with high specificity and sensitivity.

In immunodiagnostic technologies which involve mammalian sera, chicken immunoglobulin (IgY) is considered the best antibody. This antibody does not cross react with mammalian IgG due to immunological differences, there is no cross-reactivity with Rh factor, is cost effective and easy to produce, making them desirable for research and diagnostics (Bird and Thorpe, 2009). Furthermore, IgY is produced in the chicken egg yolk therefore does not require bleeding of animals, is produced rapidly in large amounts and are stable for long periods at 4°C. Additionally, chickens can produce specific IgY against small amounts of antigens with poor immunogenicity (Amro *et al.*, 2018). In the search of reliable serological diagnostic tools, it is imperative to find antibodies of highest specificity and their likelihood of misdiagnosis. Therefore, monoclonal, highly specific antibodies such as peptide antibodies and single chain variable domain fragments are preferred.

Chicken immunoglobulin genes are also used to create a large library of antibody fragments such as single chain variable fragments (scFvs) and antigen binding fragments (Fabs) for phage display technology. The scFvs are displayed as fusion coat proteins on bacteriophages and recombinantly expressed in bacterial cells. They can recognise a variety of antigens and through several rounds of selection against a target antigen, mono-specific scFvs can be isolated, recombinantly produced and used for diagnostic and research purposes (Van Wyngaardt *et al.*, 2004).

This chapter reports the production and purification of chicken IgY against the recombinant *T. b. brucei* cathepsin L-like cysteine protease (*Tbb*CATL). The cross-reactivity of anti-*Tbb*CATL IgY, anti- N-terminal *Tco*CATL peptide and anti-*Tvi*CATL antibodies was evaluated by western blot and reported. Nkuku® phage library was used to pan for scFvs against recombinant *Tvi*CATL and the results are also reported in this chapter.

3.2 Materials and methods

3.2.1 Materials

Chicken IgY antibodies and ELISA: Freund's complete and incomplete adjuvant were purchased from Sigma-Aldrich (USA). NUNC®-Immuno™ Maxisorb 96 well plates were purchased form ThermoFisher Scientific (USA), Bovine serum albumin (BSA) and 2,2-azino-bis (3-ethylbenzothiazoline-6-sulphonic) acid (ABTS) were from Roche (Germany) Polyethylene glycol (PEG), M_r 6000, was purchased from Merck (Germany). Non-fat milk powder was purchased from VWR Life science (USA).

Affinity purification: AminoLink® Coupling Resin was purchased from ThermoScientific (USA).

Western blot analysis: Pre-stained marker was purchased from Separation Scientific (USA) and BioTrace™ nitrocellulose was from PALL Life Sciences (USA).

Antibodies: Chicken anti- N-terminal *Tco*CATL peptide IgY and Chicken anti-*Tvi*CATL catalytic domain were previously produced in-house and rabbit anti-chicken IgY horseradish peroxidase (HRPO) antibodies were from Sigma-Aldrich (USA).

Phage display technology: *E. coli* TG1 cells, Nkuku® phagemid library and M13KO7 wild type helper phages were obtained from Prof. Dion du Plessis and Dr Jeanni Fehrsen (Immunology Division, Onderstepoort Veterinary Institute, SA). NUNC® Immuno™ tubes and triethylamine were purchased from ThermoFisher Scientific (USA). Indolyl-β-D-galactopyronoside (IPTG) was purchased from Fermentas (Vilnius, Lithuania). Ampicillin was purchased from Amresco (USA) and kanamycin from Life Technologies (UK). Sterile 0.22 µM filters and 96 well microtitre plates were purchased from PALL Life Sciences (USA) and Nest Bio Tech co (China) respectively.

Antibodies for phage ELISA: Goat anti-mouse HRPO conjugate antibodies were purchased from Jackson ImmunoResearch Inc (USA), mouse anti-M13 antibodies were purchased from ThermoFisher Scientific (USA) and goat anti-rabbit HRPO conjugate was purchased from Jackson ImmunoResearch Inc (USA).

3.2.2 Raising anti-*Tbb*CATL antibodies in chickens

Two chickens were immunised with purified GST-tagged *Tbb*CATL by intramuscular injection in the breast muscle. The first immunisation was done with an emulsion of 50 µg of GST-*Tbb*CATL per chicken in 500 µl PBS with an equal volume of Freund's complete adjuvant, followed by three booster immunisations prepared in the same manner with Freund's

incomplete adjuvant every two weeks. Eggs were collected throughout the immunisation period.

3.2.3 Isolation of chicken anti-*Tbb*CATL IgY

Antibodies were isolated from egg yolks as per Polson *et al.* (1985) and Goldring and Coetzer (2003). The egg yolks were separated from the albumin and rinsed with distilled water. The yolk volume was measured and mixed with double the volume of IgY buffer [100 mM Na-phosphate buffer, pH 7.4 containing 0.02% (w/v) NaN$_3$] in a beaker. To this mixture, 3.5% (w/v) polyethylene glycol (PEG) M$_r$ 6000 was added and dissolved by gentle stirring, followed by centrifugation (4420 x g, RT, 30 min). The supernatant was filtered through cotton wool, the volume of the filtrate was measured and to that, 8.5% (w/v) PEG 6000 was added and dissolved by gentle stirring followed by centrifugation (4420 x g, RT, 10 min). The resulting pellet was dissolved in IgY buffer using a volume equal to the original egg yolk volume and 12% (w/v) PEG 6000 dissolved and the precipitated IgY collected by centrifugation (12 000 x g, RT, 10 min). The resulting pellet was dissolved in 1/6 original egg yolk volume and the concentration determined: [IgY] = A$_{280}$ ÷ 1.25 (extinction coefficient; Goldring and Coetzer, 2003).

3.2.4 Evaluation of anti-GST-*Tbb*CATL IgY by enzyme linked immunosorbent assay (ELISA)

The production of antibodies during the immunisation period was evaluated by enzyme-linked immunosorbent assay (ELISA). A 96 well NUNC-Immuno™ Maxisorp ELISA plate was coated with 150 µl of 1 µg/ml GST-tagged *Tbb*CATL and incubated at 4ºC for 16 hours. The unoccupied spaces in the wells were blocked with 200 µl of 0.5% (w/v) BSA-PBS for 1 hour at 37ºC. The wells were washed with 3 x 300 µl PBS-Tween-20 [PBS with 0.1% (v/v) Tween-20] using the Biotek ELx50 plate washer (BioTek® Instruments Inc., USA). The isolated anti-*Tbb*CATL IgY antibodies were added (100 µl of 100 µg/ml) to the wells and incubated at 37ºC for 2 hours followed by washing step and incubation with rabbit anti-chicken antibodies for 1 hour. The wells were washed as previously followed by incubation with substrate solution [0.05% (w/v) ABTS. 0.005% (v/v) H$_2$O$_2$ in 150 mM citrate phosphate buffer, pH 5.5] for 1 hour. The absorbance readings were measured at 405 nm using the VERSA max Microplate Reader (Molecular Devices, USA).

3.2.5 Affinity purification of anti-GST antibodies and evaluation by ELISA

To purify anti-GST antibodies from anti-*Tbb*CATL antibodies, glutathione (GST) was coupled to AminoLink® Coupling Resin according to the Thermo Scientific instruction manual. The resin (1 ml) was equilibrated with 6 ml of coupling buffer (0.1 M Na$_2$PO$_4$ buffer, pH 7.4, 0.1 M NaCl)

in a 10 ml BioRad column. The GST protein was diluted 4-fold in coupling buffer and added to the resin. Sodium cyanoborohydride solution (5 M NaCNBH$_3$ in 1 M NaOH) (4 µl) was added to the column and incubated overnight at 4°C on an end-over-end rotator. The flow through was collected and the column washed with 4 ml coupling buffer. The unoccupied sites were blocked with 4 ml quenching buffer (1 M Tris-HCl, pH 7.4) followed by incubation with quenching buffer containing 40 µl 5 M NaCNBH$_3$ at RT for 30 min on an end-over-end rotator. The column was washed with 10 ml of wash solution (1 M NaCl) and subsequently washed with 6 ml IgY buffer.

Purified chicken anti-*Tbb*CATL antibodies (isolated as per Section 3.2.3) were incubated with the GST affinity resin overnight at 4°C on an end-over-end-rotator. The incubation mixture was transferred into a 10 ml BioRad column and the flow through (anti-*Tbb*CATL antibodies) collected. The column was washed with IgY buffer until an A$_{280}$ of 0.02 was reached followed by elution with 20 ml of elution buffer (0.1 M glycine, pH 2.8). The elution fractions were collected, neutralised with 100 µl of quenching buffer (1 M Tris-HCl, pH 7.4) and absorbance was measured at A$_{280}$.

The purified anti-GST and anti-*Tbb*CATL were evaluated in an ELISA as described in Section 3.2.4 with the following amendments: the ELISA plate was coated with 150 µl of 1 µg/ml GST-*Tbb*CATL or GST. The isolated anti-*Tbb*CATL IgY and anti-GST IgY antibodies were added to the respected antigens (100 µl of 100 µg/ml) to the wells before incubation at 37°C.

3.2.6 Evaluation of cross-reactivity by western blot between anti-*Tbb*CATL, anti-peptide *Tco*CATL and anti-peptide *Tvi*CATL and their respective antigens

Trypanosomal cysteine proteases *Tco*CATL, *Tvi*CATL and *Tbb*CATL (20 µg) were separated by 12% reducing SDS-PAGE (Section 2.2.9) and a western blot was done as described in Section 2.2.10 with the following amendments: anti-N-terminal peptide *Tco*CATL, anti-*Tvi*CATL and anti-*Tbb*CATL IgY were the primary antibodies (10 µg/ml) and 1:10 000 of rabbit anti-chicken IgG-HRPO conjugate as the secondary antibody. The membranes were incubated in substrate solution [0.06% (w/v) 4-chloro-1-naphthol, 0.1% (w/v) methanol, 0.0015% (v/v) H$_2$O$_2$ in TBS] in the dark for 3 min and the image captured using the Syngene G: BOX (Vacutec, USA).

3.2.7 Culturing of *E. coli* TG1s for phage display

Glycerol stocks of *E. coli* TG1s cells were streaked onto TYE agar medium [1% tryptone, 0.5% yeast extract, 0.8% NaCl, 1.5% (w/v) agar] without antibiotic and incubated at 37°C overnight. A single colony was inoculated into 5 ml of 2xYT medium [1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl] and incubated at 37°C with shaking (200 rpm) for 16 hours. Another

colony was incubated in 5 ml of 2xYT medium containing 5 µg/ml of ampicillin to serve as a control in which the absence of growth indicated that the cells were not infected. A total of 450 ml of 2xYT medium was added to the 5 ml culture (1:100 dilution), grown until the OD_{600} was 0.5 (log phase) and stored at 4ºC until further use (used within 1 week).

3.2.8 Titering of M13KO7 helper phages

A glycerol stock of M13KO7 helper phages was serially diluted from 10^{-2} to 10^{-13} in 2xYT medium, in sterile microfuge tubes. A total of 100 µl of each dilution was combined with 100 µl of log phase *E. coli* TG1 cells and incubated at RT for 5 min. The mixture was combined with 4 ml of warm 0.7% (w/v) agar, plated onto pre-warmed 2xYT agar medium without antibiotic and left to solidify at RT followed by incubation at 37ºC for 16 hours. The titre of the stock was determined by calculating the plaque forming units (pfu) as follows:

$$\text{pfu/ml} = \frac{\text{Number of plaques x serial dilution}}{0.5 \times 0.1 \text{ ml}}$$

3.2.9 Culturing of M13KO7 helper phage

A single plaque (from 10^{-10} dilution) was aseptically punched out of the agar plate using a sterile tip and inoculated into 4 ml of 2xYT medium containing 40 µl of the log phase *E. coli* TG1 cells and incubated at 37ºC with shaking (100 rpm) for 2 hours. After 2 hours, 2.2 ml of the culture was diluted into 400 ml of 2xYT medium in a 2 L baffled flask and incubated at 37ºC with shaking (100 rpm) for 1 hour followed by addition of 50 mg/ml of kanamycin and incubated overnight (conditions were kept the same). The culture was centrifuged (10 800 x *g,* 4ºC, 15 min), 20% (w/v) PEG 6000 in 2.5 M NaCl was added to ¼ volume of supernatant (200 ml) and incubated on ice for 30 min followed by centrifugation (10 800 x *g,* 4ºC, 15 min and 2 000 x *g*, 4ºC, 2 min). The resulting pellet was resuspended in 6 ml of PBS and filter sterilised using a 0.22 µM filter. The phages were titred as previously (Section 3.2.8) and the concentration was determined by calculating the pfu/ml from the resulting plaques.

3.2.10 Culturing of the Nkuku® phagemid library

Using the phagemid glycerol stock, 250 µl was inoculated into 500 ml of 2xYT medium containing 100 µg/ml ampicillin and 2% (w/v) glucose in a 2 L baffled flask and incubated at 37ºC with shaking (240 rpm) until an OD_{600} of 0.5 was reached. A total of 100 ml was removed from the culture and added to a 250 ml flask. The rest of the culture was incubated for a further 3 hours and centrifuged (3 300 x *g*, 4ºC, 15 min). The resulting pellet was resuspended in 1 ml 2xYT medium and stored in 15% (v/v) glycerol at -80ºC.

M13KO7 helper phage (8×10^9 pfu/ml) was added to the 100 ml culture and incubated at 37ºC, 30 min, no shaking, followed by 30 min with shaking (100 rpm). The culture was centrifuged (3 300 x *g*, at 4ºC, 15 min) and the pellet was resuspended in 1 L 2xYT medium containing 100 µg/ml ampicillin, and 25 µg/ml kanamycin in a 5 L flask. This was incubated at 30ºC with shaking (240 rpm) for 16 hours. The culture was centrifuged (3 300 x *g*, 4º, 15 min). To ¼ of the supernatant (200 ml), 20% (w/v) PEG 6000 in 2.5 M NaCl was added (in centrifuge tubes) and incubated on ice for 1 hour followed by centrifugation (3 300 x *g*, 4ºC, 15 min). The pellet was resuspended in 20 ml PBS and centrifuged (11 000 x *g*, 4ºC, 2 min). The resulting supernatant was filter-sterilised into a sterile tube using a 0.22 µM filter and stored at 4ºC for panning.

3.2.11 Titering of the Nkuku® phagemid library

A serial dilution of 10^{-2} to 10^{-7} from the Nkuku® phagemid library was done in 2xYT medium. The dilution was mixed with an equal volume of log phase *E. coli* TG1 cells (100 µl each) and incubated at RT for 5 min. The mixture was plated onto pre-warmed 2xYT agar medium containing 100 µg/ml ampicillin and 2% (w/v) glucose. To confirm that the *E. coli* TG1 cells were not infected, 100 µl of the culture was also plated and all the plates incubated at 30ºC overnight. The concentration of the Nkuku® phagemid library stock was determined by calculating the colony forming units (cfu/ml) as follows:

$$\text{cfu/ml} = \frac{\text{Number of colonies x dilution factor}}{0.5 \times 0.1}$$

3.2.12 Selection/ biopanning of scFvs against *Tvi*CATL

Two immunotubes were coated with 3.5 ml of 100 µg/ml of *Tvi*CATL and incubated overnight at 4ºC. The tubes were washed three times with PBS to remove unbound protein. The unoccupied sites were blocked with 5 ml of 2% (w/v) non-fat milk-PBS and incubated at RT for 1 hour, two control tubes were simultaneously blocked with 2% (w/v) milk-PBS. The tubes were washed twice with PBS-Tween-20 and subsequently washed twice with PBS. A volume of 3.5 ml of milk-PBS was added to the protein coated tubes and 3.5 ml of 1×10^{13} Nkuku® phagemid in 2% (w/v) milk-PBS Tween-20 was added to the control tubes and incubated at 37ºC for 30 min to eliminate milk binders from the library. The blocked Nkuku® phages were added to the protein coated immunotubes and incubated at RT for 30 min on an end-over-end rotator and 90 min standing to allow binding of high affinity phages to the immobilised protein.

The tubes were washed 20 x with PBS-Tween-20, and 20 x with PBS to remove unbound phages. A volume of 1 ml of 100 mM triethylamine (made up just before use) was added to the tubes and incubated at RT for 10 min to elute the bound phages. The eluted phages were

transferred to a 2 ml microfuge tube containing 0.5 ml Tris-HCl buffer, pH 7.4. A volume of 1 ml of the phages was added to a 50 ml falcon tube containing 5 ml of log phase *E. coli* TG1 cells and incubated at 37ºC for 30 min. The cells were centrifuged (3300 x *g*, 4ºC, 10 min) and the resulting pellet was resuspended in 1 ml 2xYT medium w/o antibiotic. A volume of 333 µl was plated onto three 2xYT agar plates containing 100 µg/ml ampicillin and 2% (w/v) glucose and incubated overnight at 30ºC.

The colonies were resuspended in 4 ml 2xYT medium of which 3 ml was stored in 15 % (v/v) glycerol at -80ºC. From the remaining 1 ml, 500 µl was added to 50 ml 2xYT medium containing 100 µg/ml ampicillin and 2% (w/v) glucose and incubated at 37ºC with shaking (200 rpm) until the OD_{600} reached 0.5. A concentration of 8 x 10^9 pfu/ml of M13KO7 was added to 5 ml of the culture and incubated at 37ºC for 30 min. The culture was centrifuged (3300 x *g*, 4ºC, 10 min, slow deceleration). The pellet was resuspended in 25 ml 2xYT medium containing 100 µg/ml ampicillin and 25 µg/ml kanamycin. The culture was incubated overnight at 30ºC with shaking (240 rpm). The culture was centrifuged (3300 x *g*, 4ºC, 20 min, slow deceleration). A volume of 5 ml 20% (w/v) PEG 6000 in 2.5 M NaCl was added to the supernatant and incubated on ice for 1.5 hours. This mixture was centrifuged (3300 x *g*, 4ºC, 30 min, slow deceleration), and the pellets were resuspended in 1 ml PBS and subsequently centrifuged (11 000 x *g*, 4ºC, 2 min). The phages were filter-sterilised with a 0.22 µM filter and either 0.5 ml (unknown concentration) or 1 x 10 ^{13} cfu/ml of the filtrates was used in the next round of panning to compare the two methods.

3.2.13 Screening of phages by polyclonal ELISA

Duplicate wells of a 96 well microtitre plate were coated with 100 µl of *Tvi*CATL (10 µg/ml and 1 µg/ml) followed by incubation overnight at 4ºC. Control wells were coated with 3% (w/v) BSA-milk. The wells were blocked with 200 µl of 3% (w/v) PBS-milk and incubated at 37ºC for 1 hour followed by incubation with 50 µl of 6% (w/v) BSA-PBS, containing 0.1% (v/v) Tween-20 and 50 µl of the screened phages diluted in PBS (1:1). Nkuku® phages were used for the control wells and incubated at 37ºC for 2 hours. The wells were washed 3 x with 295 µl PBS-0.1% (v/v) Tween-20 followed by addition of 100 µl of 100 ng/ml anti-M13 antibody (1:8000 in 3% (w/v) BSA-PBS) and incubated at 37ºC for 1 hour. The wells were washed as before and incubated with 100 µl of goat anti-mouse IgG conjugated to HPRO [1:1000 in 3% (w/v) BSA-PBS] for 1 hour at 37ºC. The wells were washed as before and incubated with 100 µl ABTS [150 mM citrate-phosphate buffer, pH 5, 0.015% (w/v) ABTS, 0.005% (v/v) H_2O_2] at 37ºC for 1 hour, and the absorbance measured at 405 nm using a VERSA max Microplate Reader (Molecular Devices, USA).

3.2.14 Screening of phages by monospecific ELISA

Phages from panning round two and three were serially diluted (10^{-5} to 10^{-10}) as described in Section 3.2.11. Single colonies were inoculated into 100 µl/well of 2xYT medium, in a sterile 96 well microtitre culture plate and incubated at 30ºC for 16 hours with shaking (220 rpm). In a new plate, 2 µl of the overnight culture was inoculated in 150 µl 2xYT medium containing 100 µg/ml ampicillin and 2% (w/v) glucose and incubated at 37ºC for 2.5 hours with shaking (220 rpm). The remaining culture was stored in 15% (v/v) glycerol at -80ºC (this was referred to as the master plate). At the end of 2.5 hours, 50 µl/well of 2 x 10^9 M13KO7 helper phages (diluted in 2xYT medium) was added and incubated at 37ºC for 30 min with no agitation. The culture was centrifuged (600 x g, 4ºC, 10 min) and the pellet was aseptically resuspended in 150 µl/well 2xYT medium containing 100 µg/ml ampicillin, 25 µg/ml glucose and incubated overnight at 30ºC with shaking (220 rpm). A non-sterile 96 well microtitre plate was coated with 100 µl of 10 µg/ml *Tvi*CATL and incubated at 4ºC for 16 hours.

The unoccupied sites of antigen coated plate were blocked with 3% (w/v) BSA-PBS (200 µl/well) and incubated at 37ºC for 1 hour. The overnight culture was centrifuged (600 x *g*, 4ºC, 10 min) and 50 µl of the supernatant was added to the blocked plate as well as 50 µl 6% (w/v) BSA-PBS containing 0.1% (v/v) Tween-20 and incubated at 37ºC for 2 hours. The plate was washed three times with 300 µl PBS-Tween-20 using the Biotek ELx50 plate washer followed by addition of 100 µl/well of mouse anti-M13 antibody (1:8000 in 3% (w/v) BSA-PBS) and incubation at 37ºC for 1 hour. The plate was washed as before and 100 µl/well of goat anti-mouse HRPO (1:1000 in 3% (w/v) BSA-PBS) added and incubated at 37ºC for 1 hour. The plate was washed as before and 100 µl of ABTS [150 mM citrate-phosphate buffer, pH 5, 0.015% (w/v) ABTS, 0.005% (v/v) H_2O_2] was added and incubated at 37ºC for 1 hour. The absorbance was measured at 405 nm using the VERSA max Microplate Reader (Molecular Devices, USA).

3.3 Results

3.3.1 Evaluation of anti-GST-*Tbb*CATL IgY by ELISA

The IgY isolated from one egg from each chicken collected each week of the immunisation period, was incubated in a GST-*Tbb*CATL coated ELISA plate. Antibody production in both chickens increased over time (Fig 3.1). Chicken 1 showed a peak of IgY production at 3 weeks and a slight decrease over weeks 5-7, before reaching another sustained peak at weeks 8-16. Chicken 2 showed a gradual increase in antibody production especially beyond 15 weeks. The A_{405} readings were generally quite low for both chickens apart form weeks 15-16 for chicken 2 (Fig. 3.1).
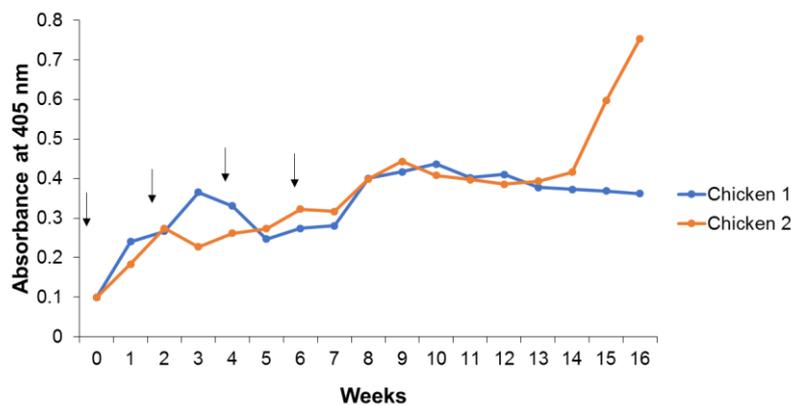
Figure 3.1: Evaluation of chicken anti-*Tbb*CATL IgY production by ELISA. A 96 well plate was coated with 1 µg/ml of *Tbb*CATL. The primary antibody was IgY isolated from single eggs from weeks 0-16 at 100 µg/ml and the secondary antibody was rabbit anti-chicken IgY antibodies conjugated to HRPO, the substrate was ABTS·H_2O_2 and the absorbance was measured at 405 nm after 1 hour and readings were corrected by subtracting the negative control. Arrows are an indication of immunisations.

3.3.2 Removal of anti-GST antibodies using a GST affinity column

The chickens were immunised with *Tbb*CATL containing a GST tag, therefore, to separate anti-GST antibodies from anti-*Tbb*CATL antibodies the total IgY was affinity purified using a GST affinity column. The A_{280} readings of the unbound fraction showed elution of the anti-*Tbb*CATL antibodies over a 50 ml volume while A_{280} readings of anti-GST antibodies showed that the antibodies were eluted from the column between 50 ml and 56 ml (Fig. 3.2).
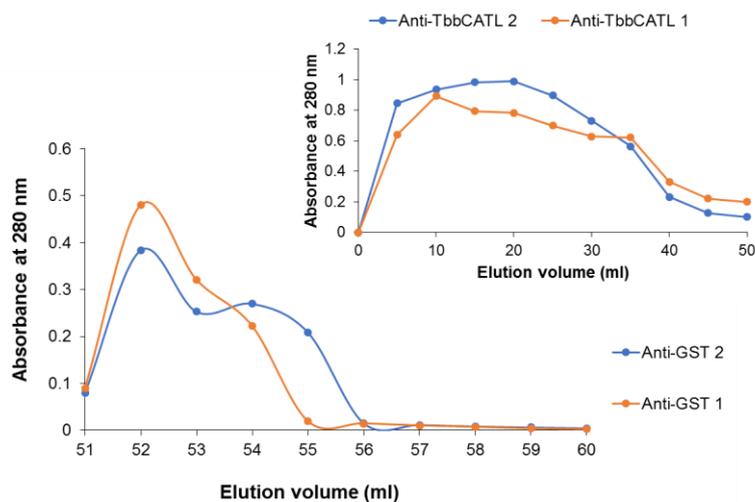
Figure 3.2: Elution profile of anti-GST antibodies from GST affinity column. IgY isolated from chickens 1 and 2 at weeks 8-11 and weeks 14-16 respectively was used to affinity purify anti-GST antibodies on a GST affinity column. The bound anti-GST antibodies were eluted with 0.1 M glycine-HCl, pH 2.8 and A_{280} of the fractions collected measured. The inset shows the A_{280} readings of the unbound fraction which contains the anti-*Tbb*CATL IgY antibodies from chickens 1 and 2.

3.3.3 Detection of GST-*Tbb*CATL by anti-*Tbb*CATL antibodies in an ELISA

Following separation from anti-GST antibodies, the detection of GST-*Tbb*CATL by the anti-*Tbb*CATL antibodies was evaluated in an ELISA. Crude IgY (before affinity purification) and pre-immune antibodies were used as controls. Antibodies from chickens 1 and 2 were able to detect GST-*Tbb*CATL with anti-*Tbb*CATL IgY showing higher absorbance values compared to anti-GST IgY (Fig. 3.3) which means that there were more antibodies against *Tbb*CATL compared to the GST tag. The detection of GST by affinity purified anti-GST IgY antibodies was also evaluated in an ELISA (Fig. 3.4). The affinity purified anti-GST antibodies from chickens 1 and 2 were able to detect GST, showing the highest absorbance readings followed by crude IgY (before affinity purification), while the anti-*Tbb*CATL antibodies remaining after removal of anti-GST antibody showed negligible detection of GST (similar to non-immune control antibodies).
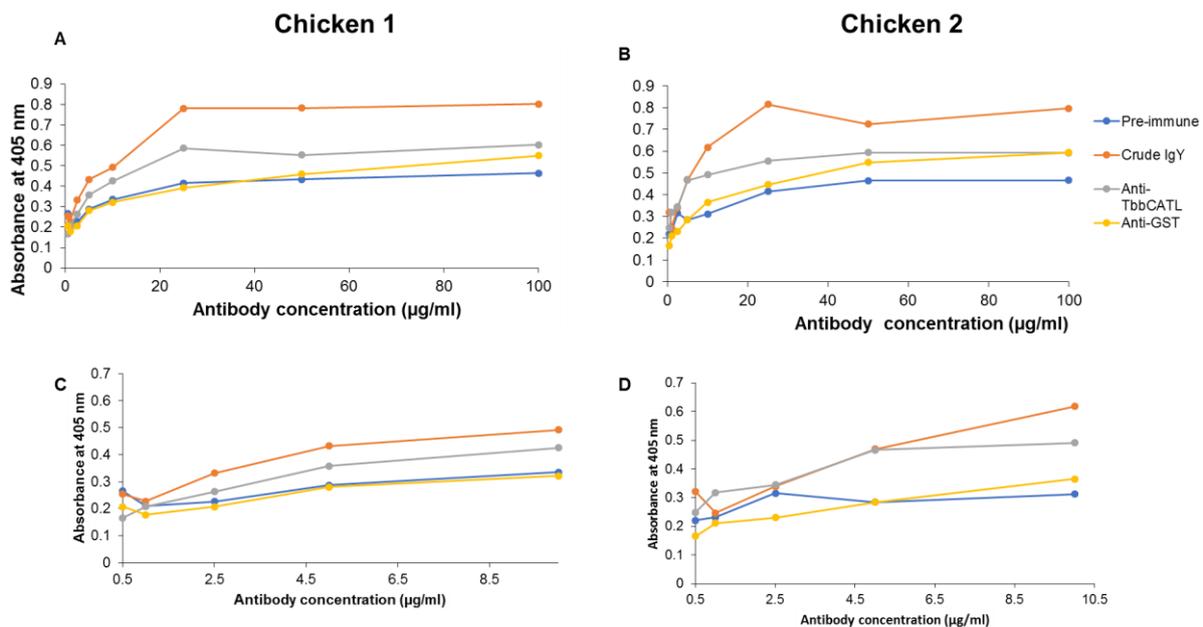
Figure 3.3: Detection of *Tbb*CATL by anti-*Tbb*CATL and anti-GST antibodies in an ELISA. Microtitre plates (96 well) were coated with 1 µg/ml of *Tbb*CATL and incubated with increasing concentrations of anti-*Tbb*CATL IgY and anti-GST IgY antibodies (0.5 µg/ml-100 µg/ml) collected after affinity purification. The secondary antibody was rabbit anti-chicken IgY IgG-HRPO conjugate and substrate ABTS·H_2O_2. Pre-immune and crude IgY antibodies were used as controls and absorbance was measured at 405 nm. Panels **C** and **D** are an enlargement of **A** and **B** respectively, showing the absorbance values for IgY concentrations 0.5-10 µg/ml for.
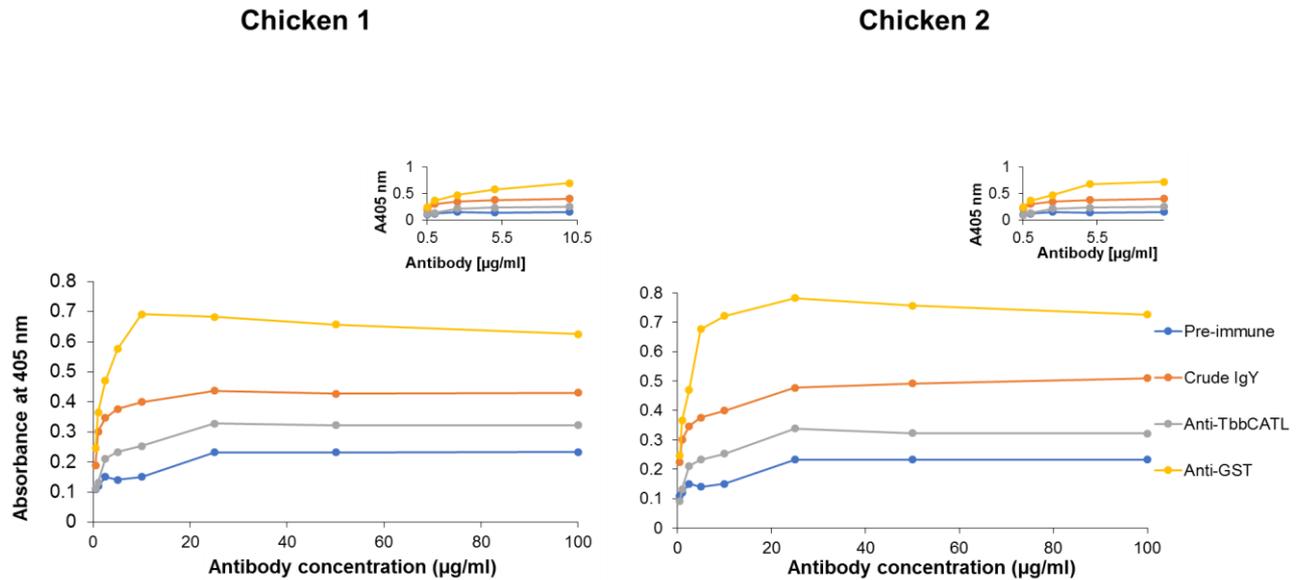
Figure 3.4: Detection of GST by affinity purified anti-GST and anti-*Tbb*CATL IgY antibodies in an ELISA. Microtitre plates were coated with 1 µg/ml of GST and incubated with increasing concentration of affinity purified chicken anti-GST IgY and anti-*Tbb*CATL IgY antibodies (0.5-100 µg/ml). The secondary antibody was rabbit anti-chicken IgG-HRPO conjugate and substrate was ABTS·H_2O_2. Pre-immune and crude IgY antibodies were used as controls and A_{405} was measured. The insets show the absorbance values for IgY concentrations 0.5-10 µg/m

3.3.4 Evaluation of cross-reactivity by western blot of anti-*Tbb*CATL, anti-*Tvi*CATL and anti-peptide *Tco*CATL antibodies and the respective antigens.

To evaluate the possibility of cross-reactivity of the antibodies, GST-*Tbb*CATL, *Tvi*CATL and *Tco*CATL, separated by reducing SDS-PAGE and transferred to a nitrocellulose membrane, were allowed to interact with anti-*Tbb*CATL, anti-*Tvi*CATL and anti-*Tco*CATL N-terminal peptide IgY in a western blot (Fig. 3.5). Anti-*Tco*CATL N-terminal peptide IgY detected the *Tco*CATL protein band at 30 kDa, *Tbb*CATL at 61 kDa and did not detect *Tvi*CATL (Fig. 3.5, panel A). Anti-*Tvi*CATL IgY antibodies detected *Tvi*CATL, with a prominent band at 29 kDa, as well as *Tbb*CATL at 61 kDa (Fig. 3.5, panel B). Anti-*Tbb*CATL IgY antibodies only detected *Tbb*CATL at 61 kDa (Fig.3.5, panel C).
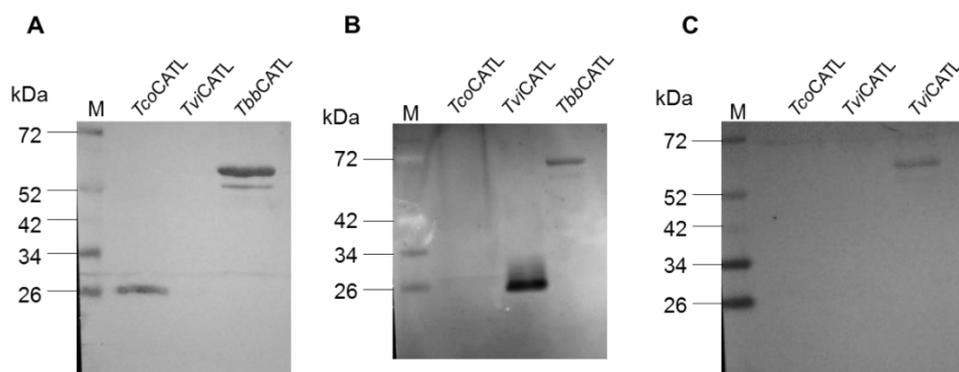
Figure 3.5: Evaluation of cross-reactivity by western blot between anti-peptide *Tco*CATL, anti-*Tvi*CATL and anti-*Tbb*CATL antibodies and their respective antigens. *Tbb*CATL, *Tvi*CATL and *Tco*CATL were separated on a 12% reducing SDS-PAGE gel and transferred onto a nitrocellulose membrane. Unoccupied sites were blocked with 5% (w/v) non-fat milk and incubated with 20 µg/ml of **A,** anti-peptide *Tco*CATL, **B**, anti-*Tvi*CATL and **C**, anti-*Tbb*CATL IgY. Rabbit-anti-chicken IgY-HRPO conjugate was the secondary antibody and 1-chloronaphthol·H_2O_2 was the substrate.

A sequence alignment was done for the cathepsin L-like cysteine proteases from *T. cruzi*, *T. vivax*, *T. b. brucei* and *T. congolense* using Clustal Omega (Chapter 2, Fig 2.1). The anti-*Tco*CATL peptide antibodies used in the western blot were against the N-terminal sequence highlighted in blue (APEAVDWRKKGAVTPVKDQGQC). The sequence identity showed that *Tbb*CATL shares 70% identity with *Tco*CATL and 62% identity with *Tvi*CATL (Table 3.1), thus explaining the cross-reactivity observed in the western blot.

Table 3.1 Sequence identities of trypanosomal cathepsin L-like cysteine proteases

	Sequence identities of CATL			
Sequence	*Tcr*CATL	*Tvi*CATL	*Tbb*CATL	*Tco*CATL
*Tcr*CATL	100%	62%	57%	55%
*Tvi*CATL	62%	100%	64%	62%
*Tbb*CATL	57%	64%	100%	70%
*Tco*CATL	55%	62%	70%	100%

3.3.5 Selection/ panning of scFvs from Nkuku® phage display library recognising *Tvi*CATL

To select for scFvs with high affinity and specificity, the Nkuku® library was panned for *Tvi*CATL binders. An unknown concentration of phages was compared to 1x 10^{13} cfu/ml for each panning round to compare the specificity of the scFvs and reproducibility of the results obtained. The scFvs were evaluated by ELISA (Fig. 3.6). At 1 x 10^{10} cfu/ml, an increase in signal was observed from pan 0 to pan 1 which is an indication of enrichment of specific scFvs (Fig. 3.6, panel A). However, this decreased in pan 2, followed by an increase in pan 3 and another decrease in pan 4 (Fig. 3.6, panel A). In the parallel experiment using an unknown

concentration of phages, enrichment is observed in pans 2 and 3, before a decrease in signal is observed in pan 4 (Fig. 3.6, panel B).
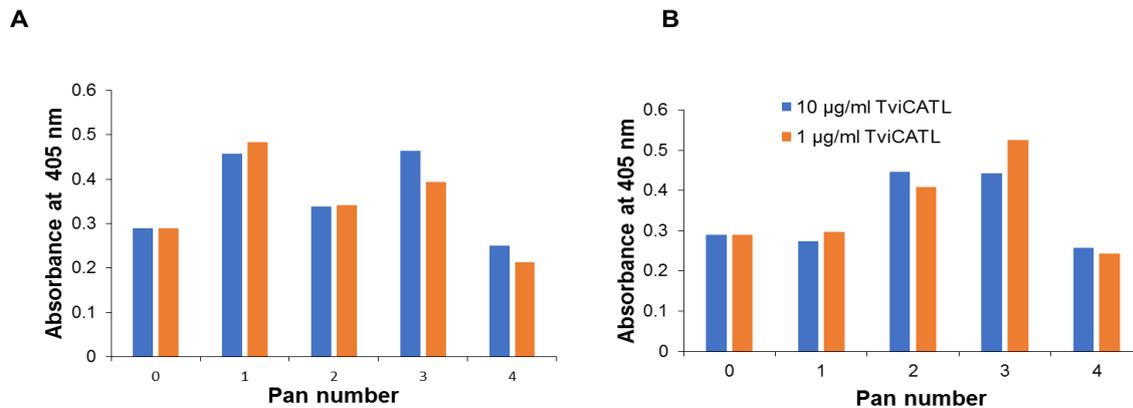
Figure 3.6: Screening of scFvs by polyclonal ELISA. The 96 well plates were coated with 10 µg/ml and 1 µg/ml *Tvi*CATL and incubated with scFvs isolated from each round of panning. Binders were detected with 1:8000 mouse anti-M13 antibody (primary) and 1:1000 goat anti-mouse-HRPO followed by ABTS·H_2O_2 substrate incubation and measuring the absorbance at 405 nm. **A**, 1 x10^{13} cfu/ml phage concentration was used for each panning round. **B**, Unknown phage concentration used in each panning round.

To select for clones with the highest specificity for *TviCATL*, the phages with the highest signals in polyclonal ELISA (Fig. 3.7) were used for screening of monospecific scFvs by ELISA. For the colonies from panning done with 1 x 10^{10} cfu/ml (Fig. 3.7, panel A), A1, A3, A11, A12 and B9 (●) gave strong signals. When the phage concentration was unknown, a strong signal was observed in H1 (which was also the strongest signal among the two experiments) and G3 (Fig. 3.7, panel B).
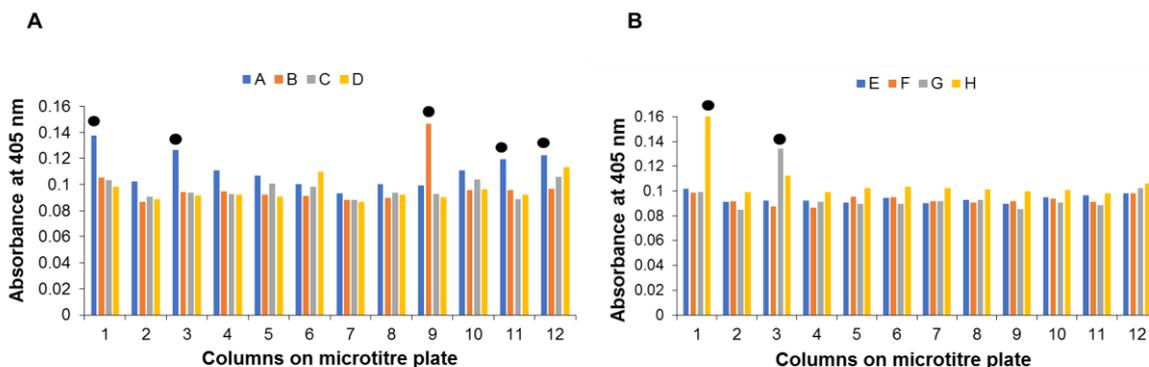
Figure 3.7: Screening of scFvs from pan 3 by monospecific ELISA. A total of 48 single colonies, infected with phages from pan 3, were cultured in a 96 well microtitre plate. The phages were rescued, isolated and allowed to interact with 10 µg/ml *Tvi*CATL. Mouse anti-M13 antibody (1:8000), goat anti-mouse HRPO conjugate antibody (1:10 000) and ATBS·H_2O_2 substrate were used. The absorbance was measures at 405 nm. **A**, 1 x10^{13} cfu/ml phage concentration used for each panning round. **B**, Unknown phage concentration used during panning.

3.4 Discussion

Antibody detection tests for AAT diagnosis are often based on the whole parasite lysate or host antibodies as a target (Torres *et al.*, 2018). This poses a challenge as the lysate preparation methods are not standardised and propagation techniques (especially for *T. vivax*) are often difficult. Additionally, species specific diagnosis of AAT has been difficult due to the high level of homology between the causative species (Boulangé *et al.*, 2017).

Lysosomal cysteine proteases from *T. b. brucei*, *T. congolense* and *T. vivax* are the major antigens in cattle infections. They are present throughout the life cycle of the parasite and highly immunogenic. There is also a correlation between cysteine protease-specific IgG response and trypanotolerance, hence the high interest in these proteases (Authié *et al.*, 2001). The results presented in this chapter reported the analysis of antibodies against recombinant *Tbb*CATL, *Tco*CATL and *Tvi*CATL as potential diagnostic tools.

There are various ways to produce antibodies with different specificities. In previous studies, anti-peptide antibodies were produced against free and carrier-conjugated peptides corresponding to a discrete amino acid sequence of *Tbb*CATL(Troeberg *et al.*, 1997) . Anti-peptide antibodies were raised in chickens and rabbits and the two forms of the peptide were found to be equally immunogenic in both species. High antibody titres were observed three weeks after initial immunisation with the conjugated peptide and after eight weeks with the free peptide. Antibodies against trypanosomal cysteine proteases have also been detected in trypanosome infected cattle (Authié *et al.*, 2001; Eyssen *et al.*, 2018). Anti-*Tco*CATL antibodies have also been raised in rabbits where it was found that the use of Freund's adjuvant elicited a strong antibody response (Huson *et al.*, 2009).

In the present study, antibodies against recombinantly expressed GST-*Tbb*CATL were produced in chickens and together with previously produced anti-*Tco*CATL N-terminal peptide and anti-*Tvi*CATL antibodies, were evaluated by western blot to determine possible cross-reactivity with the respective antigens. The results showed the presence of cross-reactivity of the anti-*Tco*CATL N-terminal peptide and anti-*Tvi*CATL antibodies with *Tbb*CATL, even though anti-peptide antibodies are known to have a specificity similar to that of monoclonal antibodies as they are directed to a specific peptide in a protein (Fig. 2.1). The observed cross-reactivity was expected since *Tco*CATL and *Tbb*CATL share 70% sequence identity (Table 3.1). The reactivity observed between anti-*Tvi*CATL antibodies and *Tbb*CATL is also an indication of sequence identity between these proteases. However, despite the broad specificity of polyclonal antibodies and high percentage identity between the three proteases, the anti-*Tbb*CATL antibodies produced here were specific for *Tbb*CATL and did not cross-react with either *Tco*CATL or *Tvi*CATL which makes them a promising tool for the specific

detection of *T. b. brucei* infections.

Due to the challenges that were encountered with purification of the recombinantly expressed GST-*Tbb*CATL fusion protein, the fusion protein was used for antibody production and the anti-GST antibodies were removed from the anti-*Tbb*CATL antibody preparation. However, it may be preferred to affinity purify the antibodies using *Tbb*CATL (after the removal of the fusion protein tag) rather than GST to ensure that none of the anti-GST antibodies are left in the antibody sample. Furthermore, further experiments will require the removal of the GST tag because it seems to interfere with applications of the protease such as activity assays as mentioned in Chapter 2 (Section 2.4). This would also ensure that the antibodies produced are only against the protein of interest and not the fusion tag.

Antibodies derived from phage display have been used for diagnostic purposes. The main challenge with commercial antibodies is that they have poor specificity, cannot recognise the target antigen or are not available. Phage display has been used for biopanning of human recombinant scFvs against a *T. cruzi* P2β ribosomal protein (Grippo *et al.*, 2011) and *Plasmodium falciparum* surface protein Pfs25, *P. falciparum* chitin 1 and *P. falciparum* circumsporozoite protein (Isaacs *et al.*, 2011).

In the present study, the Nkuku® phagemid library was used to select for scFvs with high affinity and specificity for *Tvi*CATL. The difficulties to obtain sufficient amounts of recombinant *Tbb*CATL informed the decision to work with *Tvi*CATL as a model antigen to learn how the phage display technology and optimise the system. The Nkuku® phagemid library has a highly diverse repertoire which produces antibody fragments that can recognise haptens, proteins and viruses (Van Wyngaardt *et al.*, 2004). The rescued phages were enriched after the third round of panning. Additionally, the use of a constant concentration of phages (1 x10^{13} cfu/ml) was compared to an unknown concentration (0.5 ml) during biopanning. When 0.5 ml of rescued phages was used, only two clones showed a strong signal when evaluated by monospecific ELISA. Even though more clones showed a stronger signal when 1 x 10^{13} cfu/ml was used, it cannot be concluded from this experiment alone which of the two methods is best to use. Expression and sequence analysis of these scFv would be required for their potential use as diagnostic tools.

The antibodies and scFvs produced in this study present a promising start for the development of a serological diagnostic test for AAT. The ability of the anti-*Tbb*CATL IgY antibodies to discriminate *Tco*CATL and *Tvi*CATL makes them ideal capture antibodies for an antigen-based ELISA such as a sandwich ELISA (Fig. 3.8). For species-specific antigen detection, the capture antibodies need to recognise distinct epitopes on the target antigen. Capture or detection antibodies may also be substituted with scFvs. The successful application of

nanobodies, which are variable domains of camel antibody heavy chains, as capture antibodies in a lateral flow assay for the specific detection of *T. congolense* has been reported (Torres *et al.*, 2018). It is suggested that a *Tbb*CATL-detection ELISA could have the format illustrated in Fig. 3.8. The capture antibody could be chicken anti-*Tbb*CATL IgY, that would recognise the target antigen, *Tbb*CATL, in infected cattle serum samples. An E-tagged scFv specific for *Tbb*CATL selected to recognise a different epitope on *Tbb*CATL would be detected with a HRPO-labelled rabbit anti-E-tag antibody followed by an appropriate chromogen/substrate system (Ahmad *et al.*, 2012).
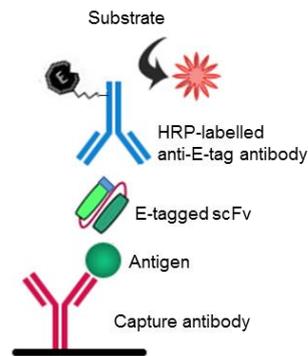
Figure 3.8: Schematic representation of a sandwich ELISA.

The results presented in this chapter showed that the anti-*Tbb*CATL antibodies produced in chickens did not cross-react with *Tco*CATL and *Tvi*CATL and are therefore promising tools for specific diagnosis of *T. b. brucei* infections. For future studies, *Tbb*CATL without a tag would be used to produce antibodies. By using *Tvi*CATL as antigen, binders were identified in the Nkuku® scFv phage display library and this can be expanded to the development of *Tbb*CATL-specific scFvs for use in diagnostic assays for *T. b. brucei* infection.

Chapter 4

General discussion

African trypanosomes are the causative agents for human African trypanosomiasis (HAT) and animal African trypanosomiasis (AAT) in sub-Saharan Africa. The human infective species are *Trypanosoma gambiense* and *T. rhodesiense* and the animal infective species are *T. b. brucei*, *T. congolense* and *T. vivax* (Muhanguzi *et al.*, 2017). These trypanosome parasites are transmitted by the tsetse fly (*Glossina* spp.), except for *T. vivax* that is also mechanically transmitted by other haematophagous flies such as horseflies (*Tabanus* spp.) and stable flies (*Stomoxys* spp.) and is consequently also found in parts of Asia and South America (Giordani

et al., 2016). The parasite escapes the host's immune response by switching genes coding for the variable surface glycoproteins, resulting in new variable antigen types. This has made it unlikely to develop a vaccine. Due to the negative impact this disease has on the economy, many research studies are aimed at controlling and treating trypanosomiasis.

Chemotherapy and chemophylaxis are the main control measures in place for trypanosomiasis, however, the use of these drugs is limited by the emergence of drug resistance and unwanted toxicity (Giordani *et al.*, 2016). The drugs used for HAT treatment depend on the causative species and stage of the disease. Suramin and pentamidine are used to treat the first stage of HAT. Suramin is effective against both HAT causing species while pentamidine is effective against gambiense HAT. Melarsoprol, nifurtimox and eflornithine are used to treat the second stage of HAT. These drugs have been associated with adverse side effects and the latter is effective against *T. b. gambiense* and not *T. b. rhodesiense* (Caffrey *et al.*, 2001; Büscher *et al.*, 2017).

The commonly used curative drug for AAT is diminazene aceturate and is used to treat infections in cattle, sheep and goats. Other drugs include isometamidine chloride, melarsomine dihydrochloride, suramin and quinapyramine which are used as prophylactic drugs. Correct administration and supportive measures are required for satisfactory treatment of ATT (Melaku and Birasa, 2013). The drugs are often administered by farmers themselves without any supervision and treatment is sometimes used without proper diagnosis (Giordani *et al.*, 2016). Due to the decrease in the development of new drugs, there is a compelling need to preserve the existing drugs.

Effective control and prevention of trypanosomiasis depend on the availability of sensitive and accurate diagnostic tools. Rapid diagnostic tests are preferred since treatment can be administered as early as possible (Torres *et al.*, 2018). Furthermore, rapid diagnostic tests are required to comply with the "ASSURED" (Affordable, Sensitive, Specific, User-friendly, Rapid and robust, Equipment free and Deliverable to end-users) criteria. Such tests have been developed for HAT diagnosis i.e. the HAT Sero-*K*-SeT (Coris BioConcept, Gembloux, Belgium) and the SD Bioline HAT 1.0 (Standard Diagnostics, Yongin, South Korea). However, these only detect *T. b. gambiense* infections and no applicable test exists for *T. b. rhodesiense* infections (Büscher *et al.*, 2017).

DNA-based assays such as PCR and LAMP are reliable, sensitive and specific for AAT diagnosis, however, are not applicable in the field as they require sophisticated equipment and experienced personnel. Alternatively, diagnosis of trypanosomiasis is often based on the whole parasite lysate antigens (derived from experimental rodent infection) to detect host antibodies against the parasite. This approach presents its own problems because a low

antibody response may result in false negative results. Sera may also contain antibodies against other co-infecting parasites such as Theileria or Babesia that may cross-react with the trypanosome lysate antigens. Antibodies are also detectable for weeks after treatment therefore it cannot be determined whether an infection is current (Torres *et al.*, 2018). This can be overcome using antigen-based tests. Recombinant antigen-based tests allow the detection of current infections as well species-specific diagnosis (Bossard *et al.*, 2010). Among the non-variant molecules, cysteine proteases have been identified and validated as potential diagnostic targets (Boulangé *et al.*, 2017; Eyssen *et al.*, 2018).

Cysteine proteases are members of the papain-like family and have been reported in several parasites including *Plasmodium* spp., *Trypanosoma* spp*.* and *Leishmania* spp. (Melaku and Birasa, 2013). In African trypanosomes, cysteine proteases are expressed throughout the life cycle and are essential for the survival of the parasite (Mbawa *et al.*, 1992). The cathepsin L-like proteases from *T. congolense*, *Tco*CATL, and *Tvi*CATL from *T. vivax* are released into the bloodstream of the infected host upon parasite lysis (Eyssen *et al.*, 2018). Additionally, trypanotolerant cattle were found to produce a high antibody response against *Tco*CATL compared to susceptible species (Authié *et al.*, 2001). Anti-*Tvi*CATL antibodies were also detected in the blood serum of *Tvi*CATL infected cattle and not in the non-infected cattle (Eyssen *et al.*, 2018) making these proteases promising targets in the development of diagnostic tools.

The overall aim of this study was to produce anti-*Tbb*CATL antibodies which can be used for specific diagnosis of *T. b. brucei* infections in antigen-detection immuno-assays. The protease *Tbb*CATL was recombinantly expressed in *E. coli* for the first time to avoid potential intellectual property restrictions associated with the *Pichia pastoris* expression system based on the strong *AOX1* promoter (Ahmad *et al.*, 2014). Since *E. coli* is quick and inexpensive to culture and high protein yields can be obtained at small scale due to the strong T7 promoter (Bill, 2014), recombinant expression of *Tbb*CATL was explored in *E. coli*.

The *TbbCATL* gene was identified (Mottram *et al.*, 1989) and amplified from the genomic DNA in a PCR reaction. The DNA fragment was ligated into PTZ57R/T TA cloning vector with no significant mutations. The gene was sub-cloned into pET32a, pET28a and pGEX-4T expression vectors and transformed into *E. coli* BL21 (DE3) cells. The homologue from *T. cruzi* has been expressed in *E. coli* (Eakin *et al.*, 1993) but previous attempts to express *Tco*CATL in a bacterial system resulted in an incorrectly folded and inactive enzyme (Boulangé *et al.*, 2001). However, since the aim of this project was to produce an antigen for antibody production, catalytic activity of *Tbb*CATL was not a requirement. An advantage of using a pET vector such as pET28a, is that the recombinant protein contains a small His-tag

that is useful for purification. A large fusion protein may be immunogenic and hence needs to be removed after purification before antibody production. A study by Lanfranco *et al.* (2008) reported the expression of the truncated form of a cathepsin B-like cysteine protease from *Leishmania braziliensis* in *E. coli* in a pQE80L expression vector. Although human cathepsin B has been expressed in *E. coli* using a pET vector (Chan and Fong, 1988), there was no *Tbb*CATL expression in either pET28a or pET32a in the present study. This may explain why there are no reports on successful expression of the trypanosomal cathepsin L-like cysteine proteases in *E. coli* using the pET vectors. However, *Tbb*CATL was expressed in soluble form using the pGEX-4T vector (which adds a 26 kDa GST tag) and purified using a glutathione-agarose resin.

In a previous study, the methylotrophic yeast *P. pastoris* was investigated as an expression system and reported suitable for the expression of *Tco*CATL (Boulangé *et al.*, 2011). Since then, recombinant cathepsin B-like and cathepsin L-like cysteine proteases from *T. b. rhodesiense* as well as *Tvi*CATL have been expressed in *P. pastoris* (Mendoza-Palomares *et al.*, 2008; Steverding *et al.*, 2012; Eyssen *et al.*, 2018). In the present study, *Tco*CATL and *Tvi*CATL were recombinantly expressed in *P. pastoris* and purified by three-phase partitioning (TPP) followed by molecular exclusion chromatography (MEC). Both *Tco*CATL and *Tvi*CATL hydrolysed the synthetic substrate Z-Phe-Arg-AMC and were inhibited by the class specific inhibitor E-64. Although *E. coli* remains the most common host for recombinant expression in both commercial and research laboratories (Rosano and Ceccarelli, 2014), it has been shown that *P. pastoris* has gained recognition for its ability to produce more challenging protein targets. Other advantages of using yeast include high protein yields, secretion of protein into the culture medium and post translation modification (Bill, 2014). A study by Shen *et al.* (2016) reported the use a non-methanol carbon source, dihydroacetone, to overcome the limitations associated with methanol. The methanol induced promoter AOX1 has also been used for recombinant expression of a phytase reporter protein without the addition of methanol but by co-expression with a positively acting transcription factor Pmlp (Ahmad *et al.*, 2014). While there is no certainty which host will produce functional yields, it has been suggested that yeast be considered in the early strategies of recombinant expression (Bill, 2014). Therefore, expression of *Tbb*CATL in *P. pastoris* should be considered for future studies.

In a previous study, Z-Phe-Arg-AMC was found to be an appropriate substrate for the native *Tbb*CATL (Troeberg *et al.*, 1996). Failure to cleave the substrate observed in the present study was accounted for by the presence of the GST tag. All attempts to cleave off the tag were unsuccessful therefore further studies were conducted with the GST-tagged *Tbb*CATL. It has been reported that the GST tag does interfere with the activity and the structure of the protein therefore like other fusion tags, should be removed prior to downstream applications of the

protein (Rosano and Ceccarelli, 2014).

As indicated before, antigen-based immuno-detection has advantages over antibody-based techniques for accurate diagnosis of AAT but relies on good antibodies. In the present study polyclonal antibodies were produced against *Tbb*CATL and producing scFvs specific for trypanosomal cathepsin L-like proteases using phage display technology was explored using *Tvi*CATL for biopanning. In a sandwich ELISA, these capturing antibodies would interact with the antigen prior to the detection antibodies, therefore must be able to recognise specific epitopes on the target antigen and be able to outcompete the host's antibodies (Torres *et al.*, 2018). In a study by Troeberg *et al.* (1997), antibodies were produced in rabbits and chickens against the native *Tbb*CATL and also against peptides corresponding to discrete sequences of *Tbb*CATL. In the present study, anti-*Tbb*CATL antibodies were produced in chickens because they are easy to handle and do not require bleeding since the antibodies are packaged in the egg yolk. Additionally, chicken IgY has been reported to be the best antibody for research diagnostics because it does not cross-react with mammalian IgG due to immunological differences and is easy to produce in a cost effective manner (Bird and Thorpe, 2009). In this study, anti-*Tbb*CATL IgY was able to discriminate *Tco*CATL and *Tvi*CATL while anti-N-terminal peptide *Tco*CATL and anti-*Tvi*CATL cross-reacted with *Tbb*CATL in a western blot. This was an interesting observation because the *Tbb*CATL antibodies were produced against the whole protein, which means they are more likely to recognise multiple epitopes leading to cross-reactivity with *Tco*CATL and *Tvi*CATL as these proteases share similarities in their sequences (Table 3.1). On the other hand, the anti-peptide *Tco*CATL antibodies are against a single specific epitope and therefore less likely to cross-react.

The cross-reactivity observed with *Tbb*CATL indicated a challenge with specific diagnosis, especially in cases where mixed infections are present. Anti-peptide antibodies are more specific because they are directed against a specific epitope on the protein (Troeberg *et al.*, 1997). In this study, anti-peptide *Tco*CATL antibodies did not cross-react with *Tvi*CATL and anti-*Tvi*CATL antibodies did not cross-react with *Tco*CATL which is promising for AAT diagnosis. Further cross-reactivity studies are required where anti-peptide *Tbb*CATL antibodies will be used. However, the polyclonal antibodies produced in this study are promising tools for the development of an ELISA based serological diagnostic test. The antibody specificity observed in this study indicated the possibility of a specific antibody-antigen complex which can be translated into a lateral flow assay (LFA).

Lateral flow assays provide a combination of biorecognition and chromatography in different formats such as sandwich ELISA. Briefly, the antibody is immobilised on a conjugate pad and a sample containing the antigen is applied to the pad and moves to other parts of the strip.

Upon migration, the antigen is captured at the conjugate pad, forming a specific antigen-antibody complex which is detected by another labelled antibody at the test line, forming a sandwich (Fig. 3.8). The excess labelled antibody is captured at the control line (Ngom *et al.*, 2010). The LFA has been described in the detection of *T. vivax* infections in cattle, based on immunodominant two invariant surface glycoproteins. The IgG from infected calves were used as capture antigens and were able to distinguish between infected and non-infected cattle, as well as *T. vivax* and *T. congolense* infected sera (Fleming *et al.*, 2016). This technique has also been reported for *T. b. gambiense* infections in humans, where IgG was directed against ISG65 (Sullivan *et al.*, 2013).

Recombinant antibodies derived from phage display technology have been applied in different diagnostic platforms such as lateral flow strips and bead-based assays. The main advantage of using recombinant antibodies is that they can be modified (Bahara *et al.*, 2013), for example, the display of antibody fragments with binding activity on the surface of bacteriophages (Kuhn *et al.*, 2016). The use of scFvs derived from monoclonal antibodies for the diagnosis of human immunodeficiency virus (HIV) has been described (Lilley *et al.*, 1994). Single chain variable fragments have also been used to control malaria infections by binding and neutralising plasmodium merozoite surface protein 1 (MSP1) and the apical membrane protein 1 (AMA1). Additionally, a peptide ligand derived from phage display was found to inhibit the invasion of *P. falciparum* into human erythrocytes (Li *et al.*, 2002; Goulart *et al.*, 2017). The use of scFvs from a buffalo serum derived library to detect a *Schistosoma japonicum* antigen in an ELISA was also reported (Hosking *et al.*, 2015).

In this study, scFvs specific for *Tvi*CATL were selected from a Nkuku® phage display library. The phages were enriched after the third round of panning and a total of seven clones gave high signals when analysed by monospecific ELISA. Further work to express and purify these scFvs needs to be done. Histidine tagged scFvs with strong affinity for target proteins can be selected and amplified, followed by purification using immobilised metal affinity chromatography. The scFvs are primarily screened by SDS-PAGE, western blot or ELISA for their ability to recognise the target antigen. Further applications of the scFvs would require expression in *E. coli* and a sequence analysis (Goulart *et al.*, 2017).

In conclusion, *Tbb*CATL was recombinantly expressed in *E. coli* and used as an immunogen for antibody production in chickens. The anti-*Tbb*CATL antibodies produced were able to discriminate *Tco*CATL and *Tvi*CATL in a western blot and together with scFvs produced using the methodology optimised here are promising leads for designing species-specific diagnostic tests for animal African trypanosomiasis.

References

Abdulla, M.-H., O'brien, T., Mackey, Z. B., Sajid, M., Grab, D. J. & Mckerrow, J. H. (2008). RNA interference of *Trypanosoma brucei* cathepsin B and L affects disease progression in a mouse model. *PLoS Neglected Tropical Diseases, 2,* 298-371.

Abras, A., Ballart, C., Llovet, T., Roig, C., Gutiérrez, C., Tebar, S., Berenguer, P., Pinazo, M.-J., Posada, E. & Gascón, J. (2018). Introducing automation to the molecular diagnosis of *Trypanosoma cruzi* infection: A comparative study of sample treatments, DNA extraction methods and real-time PCR assays. *PLoS One, 13,* 195-738.

Ahmad, M., Hirz, M., Pichler, H. & Schwab, H. (2014). Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Applied Microbiology and Biotechnology, 98,* 5301-5317.

Ahmad, Z. A., Yeap, S. K., Ali, A. M., Ho, W. Y., Alitheen, N. B. M. & Hamid, M. (2012). scFv antibody: principles and clinical application. *Clinical and Developmental Immunology, 2012,* 1-15

Álvarez, N., González, V. C. M. & Jiménez-Ruiz, A. (2002). Be careful with the primers when screening your clones by polymerase chain reaction. *Analytical Biochemistry, 308,* 189-191.

Amro, W. A., Al-Qaisi, W. & Al-Razem, F. (2018). Production and purification of IgY antibodies from chicken egg yolk. *Journal of Genetic Engineering and Biotechnology, 16,* 99-103.

Arap, M. A. (2005). Phage display technology: applications and innovations. *Genetics and Molecular Biology, 28,* 1-9.

Authié, E., Boulangé, A., Muteti, D., Lalmanach, G., Gauthier, F. & Musoke, A. (2001). Immunisation of cattle with cysteine proteinases of *Trypanosoma congolense*: targetting the disease rather than the parasite. *International Journal for Parasitology, 31,* 1429-1433.

Auty, H., Torr, S., Michoel, T., Jayaraman, S. & Morrison, L. (2015). Cattle trypanosomosis: the diversity of trypanosomes and implications for disease epidemiology and control. *Revue Scientifique et Technique, 34,* 587-598.

Babokhov, P., Sanyaolu, A. O., Oyibo, W. A., Fagbenro-Beyioku, A. F. & Iriemenam, N. C. (2013). A current analysis of chemotherapy strategies for the treatment of human African trypanosomiasis. *Pathogens and Global Health, 107,* 242-252.

Bahara, N. H. H., Tye, G. J., Choong, Y. S., Ong, E. B. B., Ismail, A. & Lim, T. S. (2013). Phage display antibodies for diagnostic applications. *Biologicals, 41,* 209-216.

Baral, T. N. (2010). Immunobiology of African trypanosomes: need of alternative interventions. *BioMed Research International, 2010,* 2-24.

Barrett, A. J. & Kirschke, H. (1981). Cathepsin B, cathepsin H, and cathepsin L. *In:* Lorand L. (ed.) *Methods in Enzymology.* New York: Academic Press. pp 3-919.

Berthier, D., Peylhard, M., Dayo, G.-K., Flori, L., Sylla, S., Bolly, S., Sakande, H., Chantal, I. & Thevenon, S. (2015). A comparison of phenotypic traits related to trypanotolerance in five West African cattle breeds highlights the value of shorthorn taurine breeds. *PLoS One, 10,* 126-498.

Bill, R. (2014). Playing catch-up with *Escherichia coli*: using yeast to increase success rates in recombinant protein production experiments. *Frontiers in Microbiology, 5,* 20-85.

Bird, C. R. & Thorpe, R. (2009). Purification of Immunoglobulin Y (IgY) From Chicken Eggs. *In*: Walker J.M.(ed) *The Protein Protocols Handbook.* Humana Press, Totowa, NJ: Springer. pp 1009-1011.

Bossard, G., Boulange, A., Holzmuller, P., Thévenon, S., Patrel, D. & Authie, E. (2010). Serodiagnosis of bovine trypanosomosis based on HSP70/BiP inhibition ELISA. *Veterinary Parasitology, 173,* 39-47.

Boulangé, A., Serveau, C., Brillard, M., Minet, C., Gauthier, F., Diallo, A., Lalmanach, G. & Authié, E. (2001). Functional expression of the catalytic domains of two cysteine proteinases from *Trypanosoma congolense. International Journal for Parasitology, 31,* 1435-1440.

Boulangé, A., Katende, J. & Authié, E. (2002). *Trypanosoma congolense*: Expression of a heat shock protein 70 and initial evaluation as a diagnostic antigen for bovine trypanosomosis. *Experimental Parasitology,* **100,** 6-11.

Boulangé, A., Khamadi, S. A., Pillay, D., Coetzer, T. H. & Authié, E. (2011). Production of congopain, the major cysteine protease of *Trypanosoma (Nannomonas) congolense*, in *Pichia pastoris* reveals unexpected dimerisation at physiological pH. *Protein Expression and Purification,* **75,** 95-103.

Boulangé, A., Pillay, D., Chevtzoff, C., Biteau, N., Comé De Graça, V., Rempeters, L., Theodoridis, D. & Baltz, T. (2017). Development of a rapid antibody test for point-of-care diagnosis of animal African trypanosomosis. *Veterinary Parasitology,* **233,** 32-38.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry,* **72,** 248-254.

Brun, R., Blum, J., Chappuis, F. & Burri, C. (2010). Human African trypanosomiasis. *The Lancet,* **375,** 148-159.

Büscher, P. (2002). Diagnosis of Human and Animal African Trypanosomiasis. *In:* Black SJ & Seed RJ. (eds.) *The African Trypanosomes.* Boston: Kluwer Academic Publishers. pp 51-63.

Büscher, P., Ngoyi, D. M., Kaboré, J., Lejon, V., Robays, J., Jamonneau, V., Bebronne, N., Van Der Veken, W. & Biéler, S. (2009). Improved models of mini anion exchange centrifugation technique (mAECT) and modified single centrifugation (MSC) for sleeping sickness diagnosis and staging. *PLoS Neglected Tropical Diseases,* **3,** 4-71.

Büscher, P., Gilleman, Q. & Lejon, V. (2013). Rapid diagnostic test for sleeping sickness. *New England Journal of Medicine,* **368,** 1069-1070.

Büscher, P., Cecchi, G., Jamonneau, V. & Priotto, G. (2017). Human African trypanosomiasis. *The Lancet,* **390,** 2397-2409.

Caffrey, C. R., Hansell, E., Lucas, K. D., Brinen, L. S., Hernandez, A. A., Cheng, J., Gwaltney Ii, S. L., Roush, W. R., Stierhof, Y.-D. & Bogyo, M. (2001). Active site mapping, biochemical properties and subcellular localization of rhodesain, the major cysteine protease of *Trypanosoma brucei rhodesiense. Molecular and Biochemical Parasitology,* **118,** 61-73.

Caffrey, C. R., Goupil, L., Rebello, K. M., Dalton, J. P. & Smith, D. (2018). Pentamidine uptake and resistance in pathogenic protozoa: past, present and future. *PLoS Neglected Tropical Diseases,* **12,** 232-239.

Carmen, S. & Jermutus, L. (2002). Concepts in antibody phage display. *Briefings in Functional Genomics,* **1,** 189-203.

Chan, M. M.-Y. & Fong, D. (1988). Expression of human cathepsin B protein in *Escherichia coli. FEBS Letters,* **239,** 219-222.

Chappuis, F., Loutan, L., Simarro, P., Lejon, V. & Büscher, P. (2005). Options for field diagnosis of human African trypanosomiasis. *Clinical Microbiology Reviews,* **18,** 133-146.

Chappuis, F. (2018). Oral fexinidazole for human African trypanosomiasis. *The Lancet,* **391,** 100-102.

Chung, Y.-B., Kong, Y., Joo, I.-J., Cho, S.-Y. & Kang, S.-Y. (1995). Excystment of *Paragonimus westermani metacercariae* by endogenous cysteine protease. *The Journal of Parasitology,* 137-142.

Conrad, U. & Scheller, J. (2005). Considerations on antibody-phage display methodology. *Combinatorial Chemistry & High Throughput Screening,* **8,** 117-126.

Cooper, C., Clode, P., Peacock, C. & Thompson, R. (2017). Host–parasite relationships and life histories of trypanosomes in Australia. *Advances in Parasitology.* 47-109.

Cropp, T. A. & Schultz, P. G. (2004). An expanding genetic code. *Trends in Genetics,* **20,** 625-630.

Dabo, N. & Maigari, A. (2017). Soft options for effective diagnosis of African animal trypanosomiasis: A Review. *International Journal of Medical Evaluation and Physical Report,* **2,** 1-9.

Deborggraeve, S. & Büscher, P. (2010). Molecular diagnostics for sleeping sickness: what is the benefit for the patient? *The Lancet Infectious Diseases,* **10,** 433-439.

Dennison, C. & Lovrien, R. (1997). Three phase partitioning: concentration and purification of proteins. *Protein Expression and Purification,* **11,** 149-161.

Desquesnes, M., Dargantes, A., Lai, D.-H., Lun, Z.-R., Holzmuller, P. & Jittapalapong, S. (2013). *Trypanosoma evansi* and *surra*: a review and perspectives on transmission, epidemiology and control, impact, and zoonotic aspects. *BioMed Research International,* **2013,** 1-20.

Eakin, A., Mcgrath, M., Mckerrow, J., Fletterick, R. & Craik, C. (1993). Production of crystallizable cruzain, the major cysteine protease from *Trypanosoma cruzi. Journal of Biological Chemistry,* **268,** 6115-6118.

Endemann, H. & Model, P. (1995). Lcoation of filamentous phage minor coat proteins in phage and in infected cells. *Journal of Molecular Biology,* **250,** 496-506.

Engstler, M., Pfohl, T., Herminghaus, S., Boshart, M., Wiegertjes, G., Heddergott, N. & Overath, P. (2007). Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. *Cell,* **131,** 505-515.

Eyssen, L. E.-A., Vather, P., Jackson, L., Ximba, P., Biteau, N., Baltz, T., Boulangé, A., Büscher, P. & Coetzer, T. H. (2018). Recombinant and native *Tvi*CATL from *Trypanosoma vivax*: Enzymatic characterisation and evaluation as a diagnostic target for animal African trypanosomosis. *Molecular and Biochemical Parasitology,* **223,** 50-54.

Ezeani, M., Okoro, H., Anosa, V., Onyenekwe, C., Meludu, S., Dioka, C. & Azikiwe, C. (2008). Immunodiagnosis of bovine trypanosomiasis in Anambra and Imo states, Nigeria, using enzyme-linked immunosorbent assay: zoonotic implications to human health. *Journal of Vector Borne Diseases,* **45,** 292-301.

Ferreira, L. G. & Andricopulo, A. D. (2017). Targeting cysteine proteases in trypanosomatid disease drug discovery. *Pharmacology & Therapeutics,* **180,** 49-61.

Fleming, J. R., Sastry, L., Wall, S. J., Sullivan, L. & Ferguson, M. A. (2016). Proteomic identification of immunodiagnostic antigens for *Trypanosoma vivax* infections in cattle and generation of a proof-of-concept lateral flow test diagnostic device. *PLoS Neglected Tropical Diseases,* **10,** 49-77.

Franco, J. R., Simarro, P. P., Diarra, A., Ruiz-Postigo, J. A. & Jannin, J. G. (2014). The journey towards elimination of *gambiense* human African trypanosomiasis: not far, nor easy. *Parasitology,* **141,** 748-760.

Gagaoua, M. & Hafid, K. (2015). Three phase partitioning system, an emerging non-chromatographic tool for proteolytic enzymes recovery and purification. *Biosensors Journal,* **5,** 100-134.

Georgieva, Y. & Konthur, Z. (2011). Design and screening of M13 phage display cDNA libraries. *Molecules,* **16,** 1667-1681.

Gibson, W., Backhouse, T. & Griffiths, A. (2002). The human serum resistance associated gene is ubiquitous and conserved in *Trypanosoma brucei rhodesiense* throughout East Africa. *Infection, Genetics and Evolution,* **1,** 207-214.

Giordani, F., Morrison, L. J., Rowan, T. G., De Koning, H. P. & Barrett, M. P. (2016). The animal trypanosomiases and their chemotherapy: a review. *Parasitology,* **143,** 1862-1889.

Glover, L., Hutchinson, S., Alsford, S., Mcculloch, R., Field, M. C. & Horn, D. (2013). Antigenic variation in A frican trypanosomes: the importance of chromosomal and nuclear context in VSG expression control. *Cellular Microbiology,* **15,** 1984-1993.

Goldring, J. D. & Coetzer, T. H. (2003). Isolation of chicken immunoglobulins (IgY) from egg yolk. *Biochemistry and Molecular Biology Education,* **31,** 185-187.

Goulart, L. R., Da S. Ribeiro, V. & Costa-Cruz, J. M. (2017). Anti-parasitic Antibodies from Phage Display. *In:* Lim, T. S. (ed.) *Recombinant Antibodies for Infectious Diseases.* Cham: Springer International Publishing. pp 155-171.

Grippo, V., Niborski, L. L., Gomez, K. A. & Levin, M. J. (2011). Human recombinant antibodies against *Trypanosoma cruzi* ribosomal P2β protein. *Parasitology,* **138,** 736-747.

Hanotte, O., Bradley, D. G., Ochieng, J. W., Verjee, Y., Hill, E. W. & Rege, J. E. O. (2002). African pastoralism: genetic imprints of origins and migrations. *Science,* **296,** 336-339.

Hoare, C. A. (1972). The trypanosomes of mammals. A zoological monograph. *A Zoological Monograph.,* **68,** 1972-2721.

Holt, H., Selby, R., Mumba, C., Napier, G. & Guitian, J. (2016). Assessment of animal African trypanosomiasis (AAT) vulnerability in cattle-owning communities of sub-Saharan Africa. *Parasites and Vectors,* **9,** 53-60.

Hoogenboom, H. R., Griffiths, A. D., Johnson, K. S., Chiswell, D. J., Hudson, P. & Winter, G. (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Research,* **19,** 4133-4137.

Hoogenboom, H. R. & Winter, G. (1992). By-passing immunisation: human antibodies from synthetic repertoires of germline VH gene segments rearranged in vitro. *Journal of Molecular Biology,* **227,** 381-388.

Hoogenboom, H. R., De Bruïne, A. P., Hufton, S. E., Hoet, R. M., Arends, J.-W. & Roovers, R. C. (1998). Antibody phage display technology and its applications. *Immunotechnology,* **4,** 1-20.

Hosking, C. G., Mcwilliam, H. E., Driguez, P., Piedrafita, D., Li, Y., Mcmanus, D. P., Ilag, L. L., Meeusen, E. N. & De Veer, M. J. (2015). Generation of a novel bacteriophage library displaying scFv antibody fragments from the natural buffalo host to identify antigens from adult *Schistosoma japonicum* for diagnostic development. *PLoS Neglected Tropical Diseases,* **9,** 42-80.

Huson, L. E. J., Authié, E., Boulangé, A. F., Goldring, J. P. D. & Coetzer, T. H. T. (2009). Modulation of the immunogenicity of the *Trypanosoma congolense* cysteine protease, congopain, through complexation with α_2-macroglobulin. *Veterinary Research,* **40,** 1-12.

Iba, Y. & Kurosawa, Y. (1997). Comparison of strategies for the construction of libraries of artificial antibodies. *Immunology and Cell Biology,* **75,** 2-17.

Isaacs, A. T., Li, F., Jasinskiene, N., Chen, X., Nirmala, X., Marinotti, O., Vinetz, J. M. & James, A. A. (2011). Engineered resistance to *Plasmodium falciparum* development in transgenic *Anopheles stephensi. PLoS Pathogens,* **7,** 100-201.

Jamonneau, V., Camara, O., Ilboudo, H., Peylhard, M., Koffi, M., Sakande, H., N'dri, L., Sanou, D., Dama, E. & Camara, M. (2015). Accuracy of individual rapid tests for serodiagnosis of *gambiense* sleeping sickness in West Africa. *PLoS Neglected Tropical Diseases,* **9,** 34-80.

Juturu, V. & Wu, J. C. (2018). Heterologous protein expression in *Pichia pastoris*: latest research progress and applications. *ChemBioChem,* **19,** 7-21.

Kennedy, P. G. (2013). Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). *The Lancet Neurology,* **12,** 186-194.

Kowalska, A., Kowalski, P. & Torres, M. Á. T. (2011). Chagas disease–American trypanosomiasis. *Polish Annals of Medicine,* **18,** 156-167.

Kuboki, N., Inoue, N., Sakurai, T., Di Cello, F., Grab, D. J., Suzuki, H., Sugimoto, C. & Igarashi, I. (2003). Loop-mediated isothermal amplification for detection of African trypanosomes. *Journal of Clinical Microbiology,* **41,** 5517-5524.

Kuhn, P., Führer, V., Unkauf, T., Moreira, G. M. S. G., Frenzel, A., Miethe, S. & Hust, M. (2016). Recombinant antibodies for diagnostics and therapy against pathogens and toxins generated by phage display. *Proteomics–Clinical Applications,* **10,** 922-948.

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

Lalmanach, G., Boulangé, A., Serveau, C., Lecaille, F., Scharfstein, J., Gauthier, F. & Authié, E. (2002). Congopain from *Trypanosoma congolense*: Drug Target and Vaccine Candidate. *Biological Chemistry,* **282,** 739-749.

Lanfranco, M. F., Loayza-Muro, R., Clark, D., Núñez, R., Zavaleta, A. I., Jimenez, M., Meldal, M., Coombs, G. H., Mottram, J. C. & Izidoro, M. (2008). Expression and substrate specificity of a recombinant cysteine proteinase B of *Leishmania braziliensis. Molecular and Biochemical Parasitology,* **161,** 91-100.

Lecaille, F., Kaleta, J. & Brömme, D. (2002). Human and parasitic papain-like cysteine proteases: Their role in physiology and pathology and recent developments in inhibitor design. *Chemical Reviews,* **102,** 4459-4488.

Li, F., Dluzewski, A., Coley, A. M., Thomas, A., Tilley, L., Anders, R. F. & Foley, M. (2002). Phage-displayed peptides bind to the malarial protein apical membrane antigen-1 and inhibit the merozoite invasion of host erythrocytes. *Journal of Biological Chemistry,* **277,** 50303-50310.

Lilley, G. G., Dolezal, O., Hillyard, C. J., Bernard, C. & Hudson, P. J. (1994). Recombinant single-chain antibody peptide conjugates expressed in *Escherichia coli* for the rapid diagnosis of HIV. *Journal of Immunological Methods,* **171,** 211-226.

Liu, B., Liu, Y., Motyka, S. A., Agbo, E. E. & Englund, P. T. (2005). Fellowship of the rings: the replication of kinetoplast DNA. *Trends in Parasitology,* **21,** 363-369.

Løset, G. Å., Roos, N., Bogen, B. & Sandlie, I. (2011). Expanding the versatility of phage display II: improved affinity selection of folded domains on protein VII and IX of the filamentous phage. *PLoS One,* **6,** 17-433.

Luckins, A. & Gray, A. (1978). An extravascular site of development of *Trypanosoma congolense*. *Nature,* **272,** 613-614.

Majiva, P., Hamers, R., Van Meirvenne, N. & Matthyssens, G. (1986). Evidence for genetic diversity in *Trypanosoma (Nannomonas) congolense*. *Parasitology,* **93,** 291-304.

Marco, M. & Miguel Coteron, J. (2012). Falcipain inhibition as a promising antimalarial target. *Current Topics in Medicinal Chemistry,* **12,** 408-444.

Matthews, K. R. (2005). The developmental cell biology of *Trypanosoma brucei*. *Journal of Cell Science,* **118,** 283-290.

Mattioli, R. C., Feldmann, U., Hendrickx, G., Wint, W., Jannin, J. & Slingenbergh, J. (2004). Tsetse and trypanosomiasis intervention policies supporting sustainable animal-agricultural development. *Journal of Food Agriculture & Environment,* **2,** 310-314.

Mbawa, Z. R., Gumm, I. D., Shaw, E. & Lonsadale-Eccles, J. D. (1992). Characterisation of a cysteine protease from bloodstream forms of *Trypanosoma congolense*. *The FEBS Journal,* **204,** 371-379.

Mcdermott, J. J. & Coleman, P. G. (2001). Comparing apples and oranges – model-based assessment of different tsetse-transmitted trypanosomosis control strategies. *International Journal for Parasitology,* **31,** 603-609.

Mckerrow, J. H. (1995). Cysteine proteases of parasites: A remarkable diversity of function. *Perspectives in Drug Discovery and Design,* **2,** 437-444.

Melaku, A. & Birasa, B. (2013). Drugs and drug resistance in African animal trypanosomosis: A review. *European Journal of Applied Sciences,* **5,** 84-91.

Mendoza-Palomares, C., Biteau, N., Giroud, C., Coustou, V., Coetzer, T., Authié, E., Boulangé, A. & Baltz, T. (2008). Molecular and biochemical characterization of a cathepsin B-like protease family unique to *Trypanosoma congolense*. *Eukaryotic Cell,* **7,** 684-697.

Miroux, B. & Walker, J. E. (1996). Over-production of proteins in *Escherichia coli:* mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology,* **260,** 289-298.

Mitchell, P. J. (2018). The constraining role of disease on the spread of domestic mammals in sub-Saharan Africa: A review. *Quaternary International,* **471,** 95-110.

Mottram, J. C., North, M. J., Barry, J. D. & Coombs, G. H. (1989). A cysteine proteinase cDNA from *Trypanosoma brucei* predicts an enzyme with an unusual C-terminal extension. *FEBS Letters,* **258,** 211-215.

Mugnier, M. R., Stebbins, C. E. & Papavasiliou, F. N. (2016). Masters of disguise: antigenic variation and the VSG coat in *Trypanosoma brucei*. *PLoS Pathogens,* **12,** 57-84.

Muhanguzi, D., Mugenyi, A., Bigirwa, G., Kamusiime, M., Kitibwa, A., Akurut, G. G., Ochwo, S., Amanyire, W., Okech, S. G. & Hattendorf, J. (2017). African animal trypanosomiasis as a constraint to livestock health and production in Karamoja region: a detailed qualitative and quantitative assessment. *BMC Veterinary Research,* **13,** 355-357.

Murray, M., Clifford, D., Gettinby, G., Snow, W. & Mcintyre, W. (1981). Susceptibility to African trypanosomiasis of N'Dama and Zebu cattle in an area of Glossina morsitans submorsitans challenge. *The Veterinary Record,* **109,** 503-510.

Naessens, J. (2006). Bovine trypanotolerance: A natural ability to prevent severe anaemia and haemophagocytic syndrome? *International Journal for Parasitology,* **36,** 521-528.

Nesterenko, M. V., Tilley, M. & Upton, S. J. (1994). A simple modification of Blum's silver stain method allows for 30 minite detection of proteins in polyacrylamide gels. *Journal of Biochemical and Biophysical Methods,* **28,** 239-242.

Ngaira, J., Olembo, N., Njagi, E. & Ngeranwa, J. (2005). The detection of non-RoTat 1.2 *Trypanosoma evansi. Experimental Parasitology,* **110,** 30-38.

Ngom, B., Guo, Y., Wang, X. & Bi, D. (2010). Development and application of lateral flow test strip technology for detection of infectious agents and chemical contaminants: a review. *Analytical and Bioanalytical Chemistry,* **397,** 1113-1135.

Nikolskaia, O. V., Lima, A. P. C. D. A., Kim, Y. V., Lonsdale-Eccles, J. D., Fukuma, T., Scharfstein, J. & Grab, D. J. (2006). Blood-brain barrier traversal by African trypanosomes requires calcium signaling induced by parasite cysteine protease. *The Journal of Clinical Investigation,* **116,** 2739-2747.

Nitsawang, S., Hatti-Kaul, R. & Kanasawud, P. (2006). Purification of papain from Carica papaya latex: Aqueous two-phase extraction versus two-step salt precipitation. *Enzyme and Microbial Technology,* **39,** 1103-1107.

Njiru, Z. K., Ouma, J. O., Bateta, R., Njeru, S. E., Ndungu, K., Gitonga, P. K., Guya, S. & Traub, R. (2011). Loop-mediated isothermal amplification test for *Trypanosoma vivax* based on satellite repeat DNA. *Veterinary Parasitology,* **180,** 358-362.

Noyes, H. A., Alimohammadian, M. H., Agaba, M., Brass, A., Fuchs, H., Gailus-Durner, V., Hulme, H., Iraqi, F., Kemp, S. & Rathkolb, B. (2009). Mechanisms controlling anaemia in *Trypanosoma congolense* infected mice. *PLoS One,* **4,** 51-70.

O'brien, T. C., Mackey, Z. B., Fetter, R. D., Choe, Y., O'donoghue, A. J., Zhou, M., Craik, C. S., Caffrey, C. R. & Mckerrow, J. H. (2008). A parasite cysteine protease is key to host protein degradation and iron acquisition. *Journal of Biological Chemistry,* **283,** 28934-28943.

Orenge, C. O., Munga, L., Kimwele, C. N., Kemp, S., Korol, A., Gibson, J. P., Hanotte, O. & Soller, M. (2012). Trypanotolerance in N'Dama x Boran crosses under natural trypanosome challenge: effect of test-year environment, gender, and breed composition. *BMC Genetics,* **13,** 87.

Otto, H.-H. & Schirmeister, T. (1997). Cysteine proteases and their inhibitors. *Chemical Reviews,* **97,** 133-172.

Paling, R., Moloo, S., Scott, J., Gettinby, G., Mcodimba, F. & Murray, M. (1991). Susceptibility of N'Dama and Boran cattle to sequential challenges with tsetse-transmitted clones of *Trypanosoma congolense. Parasite Immunology,* **13,** 427-445.

Parsons, M. (2004). Glycosomes: parasites and the divergence of peroxisomal purpose. *Molecular Microbiology,* **53,** 717-724.

Paschke, M. (2006). Phage display systems and their applications. *Applied Microbiology and Biotechnology,* **70,** 2-11.

Petrenko, V. A. & Vodyanoy, V. J. (2003). Phage display for detection of biological threat agents. *Journal of Microbiological Methods,* **53,** 253-262.

Pike, R. & Dennison, C. (1989). Protein fractionation by three phase partitioning (TPP) in aqueous/t-butanol mixtures. *Biotechnology and Bioengineering,* **33,** 221-228.

Pillay, D., Boulangé, A. & Coetzer, T. H. (2010). Expression, purification and characterisation of two variant cysteine peptidases from *Trypanosoma congolense* with active site substitutions. *Protein Expression and Purification,* **74,** 264-271.

Polson, A., Coetzer, T., Kruger, J., Von Maltzahn, E. & Van Der Merwe, K. (1985). Improvements in the isolation of IgY from the yolks of eggs laid by immunized hens. *Immunological Investigations,* **14,** 323-327.

Ponte-Sucre, A. (2016). An overview of *Trypanosoma brucei* infections: an intense host–parasite interaction. *Frontiers in Microbiology,* **7,** 21-26.

Radwanska, M., Vereecke, N., Deleeuw, V., Pinto, J. & Magez, S. (2018). Salivarian Trypanosomosis: A review of parasites involved, their global distribution and their interaction with the innate and adaptive mammalian host immune system. *Frontiers in Immunology,* **9,** 7-57.

Rakonjac, J. (2012). Filamentous bacteriophages: biology and applications. *Encyclopedia of Life Sciences,* 1-13.

Rami, A., Behdani, M., Yardehnavi, N., Habibi-Anbouhi, M. & Kazemi-Lomedasht, F. (2017). An overview on application of phage display technique in immunological studies. *Asian Pacific Journal of Tropical Biomedicine,* **7,** 599-602.

Reed, S., Bouvier, J., Pollack, A. S., Engel, J. C., Brown, M., Hirata, K., Que, X., Eakin, A., Hagblom, P. & Gillin, F. (1993). Cloning of a virulence factor of *Entamoeba histolytica.* Pathogenic strains possess a unique cysteine proteinase gene. *Journal of Clinical Investigation,* **91,** 15-32.

Rosano, G. L. & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology,* **5,** 172-181.

Rosenthal, P. J., Mckerrow, J., Aikawa, M., Nagasawa, H. & Leech, J. (1988). A malarial cysteine proteinase is necessary for hemoglobin degradation by *Plasmodium falciparum. Journal of Clinical Investigation,* **82,** 1560-1566.

Ruigrok, V. J., Levisson, M., Eppink, M. H., Smidt, H. & Van Der Oost, J. (2011). Alternative affinity tools: more attractive than antibodies? *Biochemical Journal,* **436,** 1-13.

Sajid, M. & Mckerrow, J. H. (2002). Cysteine proteases of parasitic organisms. *Molecular and Biochemical Parasitology,* **120,** 1-21.

Schechter, I. & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications,* **27,** 157-162.

Schuster, S., Krüger, T., Subota, I., Thusek, S., Rotureau, B., Beilhack, A. & Engstler, M. (2017). Developmental adaptations of trypanosome motility to the tsetse fly host environments unravel a multifaceted in vivo microswimmer system. *eLife,* **6,** 27-656.

Schwede, A., Macleod, O. J., Macgregor, P. & Carrington, M. (2015). How does the VSG coat of bloodstream form African trypanosomes interact with external proteins? *PLoS Pathogens,* **11,** 52-59.

Scott, J. K. & Smith, G. P. (1990). Searching for peptide ligands with an epitope library. *Science,* **249,** 386-390.

Serveau, C., Boulangé, A., Lecaille, F., Gauthier, F., Authié, E. & Lalmanach, G. (2003). Procongopain from *Trypanosoma congolense* is processed at basic pH: an unusual feature among cathepsin L-like cysteine proteases. *Biological Chemistry,* **384,** 921-927.

Shen, W., Xue, Y., Liu, Y., Kong, C., Wang, X., Huang, M., Cai, M., Zhou, X., Zhang, Y. & Zhou, M. (2016). A novel methanol-free *Pichia pastoris* system for recombinant protein expression. *Microbial Cell Factories,* **15,** 178-189.

Silvester, E., Mcwilliam, K. R. & Matthews, K. R. (2017). The Cytological Events and Molecular Control of Life Cycle Development of *Trypanosoma brucei* in the Mammalian Bloodstream. *Pathogens,* **6,** 8-29.

Simarro, P., Franco, J., Diarra, A., Postigo, J. R. & Jannin, J. (2012). Update on field use of the available drugs for the chemotherapy of human African trypanosomiasis. *Parasitology,* **139,** 842-846.

Siqueira-Neto, J. L., Debnath, A., Mccall, L.-I., Bernatchez, J. A., Ndao, M., Reed, S. L. & Rosenthal, P. J. (2018). Cysteine proteases in protozoan parasites. *PLoS Neglected Tropical Diseases,* **12,** 65-120.

Sixholo, J., Van Wyngaardt, W., Mashau, C., Frischmuth, J., Du Plessis, D. H. & Fehrsen, J. (2011). Improving the characteristics of a mycobacterial 16 kDa-specific chicken scFv. *Biologicals,* **39,** 110-116.

Smith, G. P. & Scott, J. K. (1993). Libraries of Peptides and Proteins Pisplayed on Filamentous Phage. *In:* Wu, R. (ed.) *Methods in Enzymology.* Amsterdam: Elsevier. 228-257.

Solano, P., Michel, J.-F., Lefrançois, T., De La Rocque, S., Sidibe, I., Zoungrana, A. & Cuisance, D. (1999). Polymerase chain reaction as a diagnosis tool for detecting trypanosomes in naturally infected cattle in Burkina Faso. *Veterinary Parasitology,* **86,** 95-103.

Sternberg, J. M. (2004). Human African trypanosomiasis: clinical presentation and immune response. *Parasite Immunology,* **26,** 469-476.

Stevens, J. & Gibson, W. (1999). The molecular evolution of trypanosomes. *Parasitology Today,* **15,** 432-437.

Stevens, J. & Brisse, S. (2004). Systematics of trypanosomes of medical and veterinary importance. *The Trypanosomiases,* **13,** 1-23.

Steverding, D. (2008). The history of African trypanosomiasis. *Parasites & Vectors,* **1,** 3.

Steverding, D., Sexton, D. W., Wang, X., Gehrke, S. S., Wagner, G. K. & Caffrey, C. R. (2012). *Trypanosoma brucei*: chemical evidence that cathepsin L is essential for survival and a relevant drug target. *International Journal for Parasitology,* **42,** 481-488.

Stijlemans, B., Cnops, J., Naniima, P., Vaast, A., Bockstal, V., De Baetselier, P. & Magez, S. (2015). Development of a pHrodo-based assay for the assessment of in vitro and in vivo erythrophagocytosis during experimental trypanosomosis. *PLoS Neglected Tropical Diseases,* **9,** 35-61.

Sullivan, L., Wall, S. J., Carrington, M. & Ferguson, M. A. (2013). Proteomic selection of immunodiagnostic antigens for human African trypanosomiasis and generation of a prototype lateral flow immunodiagnostic device. *PLoS Neglected Tropical Diseases,* **7,** 20-87.

Sunter, J. D. & Gull, K. (2016). The flagellum attachment zone:'the cellular ruler'of trypanosome morphology. *Trends in Parasitology,* **32,** 309-324.

Sutcliffe, O., Skellern, G., Araya, F., Cannavan, A., Sasanya, J., Dungu, B., Van Gool, F., Munstermann, S. & Mattioli, R. (2014). Animal trypanosomosis: making quality control of trypanocidal drugs possible. *Revue Scientifique et Technique,* **33,** 813-830.

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

Torres, J. E. P., Goossens, J., Ding, J., Li, Z., Lu, S., Vertommen, D., Naniima, P., Chen, R., Muyldermans, S. & Sterckx, Y. G.-J. (2018). Development of a Nanobody-based lateral flow assay to detect active *Trypanosoma congolense* infections. *Scientific Reports,* **8,** 9-19.

Troeberg, L., Pike, R. N., Morty, R. E., Berry, R. K., Coetzer, T. H. & Lonsdale-Eccles, J. D. (1996). Proteases from *Trypanosoma brucei brucei*: purification, characterisation and interactions with host regulatory molecules. *European Journal of Biochemistry,* **238,** 728-736.

Troeberg, L., Pike, R. N., Lonsdale-Eccles, J. D. & Coetzer, T. H. (1997). Production of anti-peptide antibodies against trypanopain-Tb from *Trypanosoma brucei brucei:* effects of antibodies on enzyme activity against Z-Phe-Arg-AMC. *Immunopharmacology,* **36,** 295-303.

Troeberg, L., Morty, R. E., Pike, R. N., Lonsdale-Eccles, J. D., Palmer, J. T., Mckerrow, J. H. & Coetzer, T. H. (1999). Cysteine Proteinase Inhibitors Kill Cultured Bloodstream Forms of *Trypanosoma brucei brucei. Experimental Parasitology,* **91,** 349-355.

Turk, B. & Turk, V. (2009). Lysosomes as "suicide bags" in cell death: myth or reality? *Journal of Biological Chemistry,* **284,** 21783-21787.

Turk, D., Podobnik, M., Kuhelj, R., Dolinar, M. & Turk, V. (1996). Crystal structures of human procathepsin B at 3.2 and 3.3 Å resolution reveal an interaction motif between a papain-like cysteine protease and its propeptide. *FEBS Letters,* **384,** 211-214.

Turk, D., Gunčar, G., Podobnik, M. & Turk, B. (1998). Revised definition of substrate binding sites of papain-like cysteine proteases. *Biological Chemistry,* **379,** 137-148.

Turk, V., Stoka, V., Vasiljeva, O., Renko, M., Sun, T., Turk, B. & Turk, D. (2012). Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics,* **1824,** 68-88.

Van Den Bossche, P., Chigoma, D. & Shumba, W. (2000). The decline of anti-trypanosomal antibody levels in cattle after treatment with trypanocidal drugs and in the absence of tsetse challenge. *Acta Tropica,* **77,** 263-270.

Van Wyngaardt, W., Malatji, T., Mashau, C., Fehrsen, J., Jordaan, F., Miltiadou, D. & Du Plessis, D. H. (2004). A large semi-synthetic single-chain Fv phage display library based on chicken immunoglobulin genes. *BMC Biotechnology,* **4,** 1472-6750.

Verma, S., Dixit, R. & Pandey, K. C. (2016). Cysteine Proteases: Modes of Activation and Future Prospects as Pharmacological Targets. *Frontiers in pharmacology,* **7,** 33-89.

Vreysen, M. J., Seck, M. T., Sall, B. & Bouyer, J. (2013). Tsetse flies: their biology and control using area-wide integrated pest management approaches. *Journal of Invertebrate Pathology,* **112,** 15-25.

Weber, J., Peng, H. & Rader, C. (2017). From rabbit antibody repertoires to rabbit monoclonal antibodies. *Experimental & Molecular Medicine,* **49,** 200-305.

Winter, G., Griffiths, A. D., Hawkins, R. E. & Hoogenboom, H. R. (1994). Making antibodies by phage display technology. *Annual Review of Immunology,* **12,** 433-455.

Yaro, M., Munyard, K. A., Stear, M. J. & Groth, D. M. (2016). Combatting African animal trypanosomiasis (AAT) in livestock: The potential role of trypanotolerance. *Veterinary Parasitology,* **225,** 43-52.

Zhao, A., Tohidkia, M. R., Siegel, D. L., Coukos, G. & Omidi, Y. (2016). Phage antibody display libraries: a powerful antibody discovery platform for immunotherapy. *Critical Reviews in Biotechnology,* **36,** 276-289.