MOLECULAR CHARACTERISATION OF METACASPASE 5 AND THE PRODUCTION OF OLIGOPEPTIDASE B-SPECIFIC SINGLE CHAIN VARIABLE FRAGMENT ANTIBODIES FOR POTENTIAL ANIMAL AFRICAN TRYPANOSOMOSIS CHEMOTHERAPIES AND DIAGNOSTICS

by

Lauren Elizabeth-Ann Eyssen

M.Sc

Submitted in the fulfilment of the

academic requirements for the

PhD degree

in

Biochemistry

School of Life Sciences

University of KwaZulu-Natal

Pietermaritzburg

2018

PREFACE

The experimental work described in this dissertation was carried out in the School of Life Science, University of KwaZulu-Natal, Pietermaritzburg, from July 2014 to March 2017, 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.

Lauren Eyssen

January 2018

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

Prof. Theresa H. T. Coetzer

DECLARATION – PLAGIARISM

I, Lauren Elizabeth-Ann Eyssen, 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 Reference section.

Lauren Eyssen

January 2018

DEDICATION

This thesis is dedicated to my grampa,

Edward Henry Middleborough (1939 - 2016).

AKNOWLEDGEMENTS

I would like to express my heartfelt thanks and appreciation to the following persons:

Firstly, to Prof Coetzer for the opportunity to complete this degree under your guidance, and your belief in my scientific ability.

To the lecturers in our department, Prof Goldring, Dr Hewer, and Prof Niesler for your time, your advice and encouragement.

To Charmaine Ahrens, Tanya Karalic, Natalie Jones for their tireless administration work which makes our lives as post graduate students much easier. To Pat Joubert for all the ordering of our desperately needed lab reagents and all the other small, but very necessary things you do. To Celeste Clarke, for your constant fixing and organising of all the necessary things. Along with Megan Brunkhorst, thank you all for your time and friendship.

To my fellow post graduates both past and current: each of you have, in some way or other, shaped me into the researcher that I am today with your advice, encouragement and even your presence. Thank you.

To the National Research Foundation (NRF) for their financial assistance.

Finally, to my families here, there and everywhere, a huge thank you for your never ending love and support during my many years at UKZN.

.

ABSTRACT

African trypanosomosis (AT) is a major obstacle in the establishment of agriculture and economic sustainability in Africa. Animal AT is responsible for large numbers of livestock succumbing to the tsetse transmitted kinetoplastid parasites, *Trypanosoma congolense* and *T. vivax*, and as a result, losses in further downstream sectors are experienced. Due to the ability of the trypanosomal parasites to undergo antigenic variation, vaccine candidates are highly unlikely. Peptidases have been identified as virulence factors and are the focus of the development of novel chemotherapies and diagnostics. The metacaspases (MCAs) are a prime example of a chemotherapeutic target and oligopeptidase B (OPB), that of a diagnostic target. Towards the validation of a chemotherapeutic target, recombinant expression was used to obtain an active peptidase which could be enzymatically characterised. Various inhibitors were investigated and their effect on the parasite, analysed. Current diagnostics are based on antibody detection, but an antigen detection format would be preferable as it could differentiate between active and cured infections as anti-trypanosomal antibodies can persist for years. Given the rural, resource-poor locations in the areas of AT incidence, an ideal rapid diagnostic test (RDT) would be robust, affordable, sensitive and specific and requiring minimal training, such as a dipstick test. The MCAs are cysteine peptidases which are found in all kingdoms other than the metazoa, and share a secondary structure fold and catalytic dyad with the metazoan caspases. Since the caspases play a role in apoptosis, it is thought that the MCAs may function in a similar manner. The single copy MCAs of *Trypanosoma* spp. and *Leishmania* spp. differ from the multicopy MCAs in that they possess a Pro-, Gln-, Tyr-rich C-terminal domain which is thought to mediate protein-protein interactions. The activity of the single copy MCAs from *T. cruzi* and *L. major* has been implicated in the cell cycle of the kinetoplastid parasite.

The aim of the project was to express, purify and enzymatically characterise the recombinant and native MCA5 from *T. congolense* and *T. vivax*. Using the 3D structures, solved by X-ray diffraction, of MCA2 from *T. b. brucei*, molecular docking studies were used to validate the inhibition potential of a published library of inhibitors, designed based on the, then, hypothetical structure of *Tbb*MCA2. Since the elucidation of the 3D structure of *Tbb*MCA2 by X-ray diffraction, the inhibitory power of the library of inhibitors against *Tbb*MCA2 and the MCA5s was investigated. The serine peptidase, OPB, has been shown to be released into the host bloodstream by dead and dying

parasites. The use of phage displayed scFv (single chain fragment variable) antibodies for the detection of OPB in serum from infected cattle is reported, towards the development of a RDT.

Recombinantly expressed *Tco*MCA5 was shown to autoprocess and over autoprocess when purified using nickel affinity chromatography. Mutagenesis of the catalytic dyad residues reduced the over autoprocessing and the mutated form was enzymatically active at a pH between 6 and 9. This active mutant and purified *Tco*MCA5 showed a prefence for Arg over Lys at the P_1 substrate position and were able to hydrolyse gelatin.

Possible novel inhibitors of *Tbb*MCA2 and the MCA5s of *T. congolense* and *T. vivax* were identified using a library of ligands (Berg library) based on the P_1 specificity of *Tbb*MCA2 and molecular docking. Commercial fluorogenic peptide substrates and inhibitors reported in literature for the characterisation of various MCAs, revealed interactions with the MCAs which should be taken into consideration when modifying the Berg ligands to achieve higher affinity for the MCAs.

The application of scFv antibodies, derived from the *Nkuku*® phagemid library, for the diagnosis of current AAT infections by the detection of OPB, released in the bloodstream of the infected mammalian host, was investigated. After the successful isolation and production of OPB-specific scFv, MCA-specific scFv antibodies can be pursued using the *Nkuku*® phagemid library. The resulting OPB-specific scFv identified a conserved peptide between *T. congolense* and *T. vivax* and was able to detect native OPB in a western blot format. It was predicted that the scFv interacted with OPB in such a way that it would restrict the hinge motion between the C-terminal catalytic and N-terminal regulatory domains of the enzyme and limit access to the active site pocket. The ability of scFv and rabbit-anti-OPB polyclonal antibody in an antigen detection ELISA with sera from *T. congolense* infected cattle indicated that detection of OPB fluctuated with parasitaemia.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

2xYT	2 x yeast extract, tryptone
2D	two dimensional
3D	three dimensional
AAT	animal African trypanosomosis
ABTS	2,2-azino-di-[3-ethylbenzthiazoline sulfonate]
AEBSF	4-(2-aminoethyl)benzenesulfonyl fluoride
AMC	7-amino-4-methylcoumarin
Amp	ampicillin
AMT	acetate-MES-Tris
ASSURED	affordability, sensitivity, specificity, user friendly, rapid and robust, equipment free, deliverable
AT	African trypanosomosis
BARP	brucei alanine rich protein
BCA	bicinchoninic acid
Boc	t-butyloxycarbonyl
bp	base pair
BP	3% (w/v) BSA-PBS
BSA	bovine serum albumin
CATL	cathepsin-L
CATT	card agglutination test
CDR	complementary determining regions
CSF	cerebrospinal fluid
DALYS	disability-adjusted life years
dH_2O	distilled H_2O
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
E-64	L-*trans*-epoxysuccinyl-leucylamido(4-guanidino)butane
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetra-acetic acid
ELISA	enzyme-linked immunosorbent assay
EMEM	Eagle's minimal essential medium
ES	expression sites

ESAG	expression site-associated gene
Fab	antigen binding fragment
FBS	foetal bovine serum
FMK	fluoromethylketone
g	relative centrifugal force
GMQE	global model quality estimation
GPI	glycosylphosphatidylinositol
GRESAG	related to expression site associated gene
h	hour
HAT	human African trypanosomosis
HRPO	horseradish peroxidase
HSP	heat shock protein
[I]	inhibitor concentration
IC_{50}	half maximal inhibitory concentration
IFN-γ	gamma interferon
Ig	immunoglobulin
IL	interleukin
IPTG	isopropyl-β -D-thiogalactopyranoside
ISG	invariant surface glycoprotein
ITS	internal transcribed spacer
Kan	kanamycin
k_{cat}	turnover number
kDa	kilo Dalton
K_i	inhibition constant
K_m	Michaelis constant
MCA	metacaspase
MeOSuc	3-methoxysuccinyl
MES	2-ethanesulfonic acid
min	minute
MP	2% skim milk powder-PBS
ND	not determined
NECT	eflornithine-nifurtimox combination therapy
NO	nitric oxide
OPB	oligopeptidase B
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline

PBS-T	0.1% (v/v) Tween-20-PBS
PCD	programmed cell death
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
p*I*	isoelectric point
PSG	phosphate saline glucose
RDT	rapid diagnostic test
RNA	ribonucleic acid
RT	room temperature
[S]	substrate concentration
sarkosyl	N-lauroylsarcosine sodium salt
SBDD	structure-based drug design
SBTI	soya bean trypsin inhibitor
scFv	single chain variable fragment
SDS	sodium dodecyl sulfate
SoTe	sorbitol, Tris-HCl, EDTA
SRA	serum resistance-associated gene
SOC	super optimal cataboliser
TAE	tris-acetate-EDTA
TBS	tris buffered saline
TB	terrific broth
TE	Tris-HCl, EDTA
TELT	Tris-HCl, EDTA, LiCl, Triton X-100
TL	immune trypanolysis
TLCK	N-tosyl-L-lysyl chloromethylketone
TNF	tumor necrosis factor
Tris	2-amino-2-(hydroxymethyl)-1,3-propandiol
TYE	tryptone, yeast extract
V_H	variable heavy chain
V_HH	variable heavy domain of camelid antibodies, nanobodies
v_i	steady-state velocity after addition of inhibitor
V_L	variable light chain
VLS	virtual ligand screening
V_{max}	maximum velocity
v_o	initial reaction velocity

VSG	variant surface glycoprotein
WHO	World Health Organisation
WT	wild type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
Z	benzyloxycarbonyl

CHAPTER 1

LITERATURE REVIEW

1.1 African trypanosomosis

The trypanosomal parasites which cause human and animal African trypanosomosis (HAT and AAT respectively), otherwise known as sleeping sickness in humans and nagana in cattle, are transmitted by the bite of an infected tsetse fly (*Glossina* spp.). In 2004, HAT cost the African continent 1.6 million disability-adjusted life years (DALYs) (Févre *et al.*, 2008). Together with serological screening, reservoir control and appropriate treatment, the disability-adjusted life-years (DALYs) has been lowered to 202 438 (GBD *et al.*, 2017). African trypanosomosis affects human health and has economic consequences due to its impact on agriculture resulting from the reduction in animal numbers and productivity (Hotez *et al.*, 2009; Jackson *et al.*, 2015). Since both HAT and AAT occur in rural areas, affect poor smallholders, and pose no immediate threat to wealthy nations, the disease is often considered neglected (Diall *et al.*, 2017).

Trypanosoma brucei gambiense is responsible for HAT in eastern and southern African, while *T. b. rhodesiense* causes HAT in western and central Africa (Welburn *et al.*, 2001a). The *T. b. gambiense* parasite causes chronic infections which can persist for years, whilst *T. b. rhodesiense* infections result in death within 6 to 8 months if untreated (Checchi *et al.*, 2008). It is estimated that 55 million people in an area of 1.08 million km^2 are at risk of *T. b gambiense* infections, whilst 6 million in 100 000 km^2 are at risk of *T. b. rhodesiense* infections (Franco *et al.*, 2017). The areas in which HAT infections are likely have halved in less than a decade, with fewer than 4 000 reported cases in 2014 (Franco *et al.*, 2017) (Fig. 1.1, panel A). However, due to the remote locations and poor coverage by surveillance systems of these areas, the number of reported cases may be higher.

Both *T. b. gambiense* and *T. b. rhodesiense* parasites are harboured in wild and domestic animals, which act as reservoirs (Anderson *et al.*, 2011). Domestic cattle are the major reservoir of *T. b. rhodesiense* in Uganda (Welburn *et al.*, 2001b) and the introduction of infected cattle from districts with established human *T. b. rhodesiense* infections, caused outbreaks in previously unaffected districts (Picozzi *et al.*, 2005).

A third subspecies of *T. brucei*, which is non-pathogenic in humans, is *T. b. brucei* and causes AAT. This parasite is distributed across sub-Saharan Africa and is restricted to non-human vertebrates. In addition to *T. b. brucei*, several other trypanosomal species cause infections in animals and are transmitted by insect vectors other than the tsetse fly. *Trypanosoma b. brucei, T. congolense and T. vivax,* cause Nagana in Africa, *T. evansi* cause Surra in northern Africa and Asia, infect a wide range of domestic animals as well as wild animals which act as reservoirs, whilst *T. simiae* and *T. suis* infect pigs (Uilenberg and Boyt, 1998; Lucius *et al.*, 2017). In addition, since the development of the trypanosome is restricted to the tsetse proboscis in *T. vivax*, the direct mechanical transmission by other haematophagous flies, such as tabanids and *Stomoxys*, occurs frequently (Rotureau and Van Den Abbeele, 2013). Consequently, *T. vivax* infections has spread beyond the borders of the sub-Saharan tsetse belt into northern Africa and south America (Jones and Dávila, 2001; Osório *et al.*, 2008; Gonzatti *et al*., 2014). The threat of AAT has restricted the development of agriculture due to the ineffective use of available land in tsetse-infested areas (Matthews, 2005; Jackson *et al.*, 2015; Diall *et al.*, 2017).

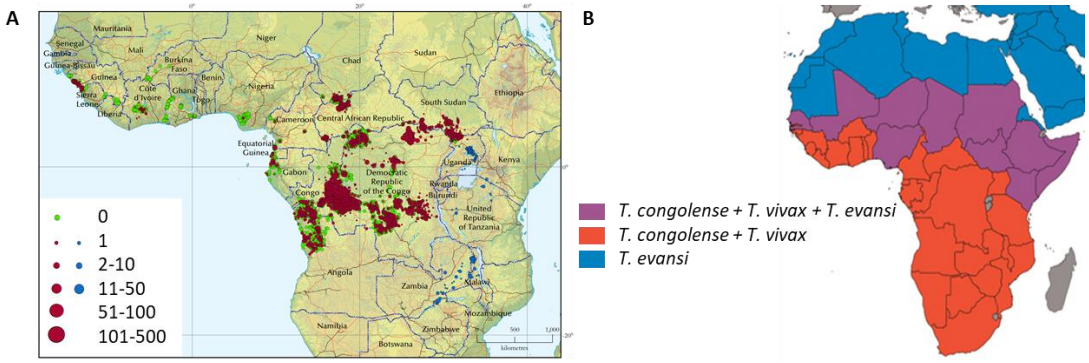


Figure 1.1: Distribution of HAT and AAT. (A) Number of *T. b. gambiense* (red) and *T. b. rhodesiense* (blue) HAT cases reported between 2010 and 2014. Areas where no cases were reported are indicated in green (Franco *et al.*, 2017). **(B)** Distribution of important animal infective trypanosomal species in Africa. The real geographical distribution of *T. congolense* and *T. vivax* in South Africa and Namibia is limited (Giordani *et al.*, 2016).

Multiple species and sub-species of animal infective trypanosomes circulate in cattle within HAT areas (Fyfe *et al.*, 2017) (Fig. 1.1, panel B). The animal infective *T. congolense,* is considered to be the most pathogenic trypanosome species in cattle (Giordani *et al.*, 2016), followed by *T. vivax* (Osório *et al.*, 2008), both of which are able to infect multiple animal species. Whilst *T. congolense* and *T. vivax* impact on both animal health and productivity across much of Uganda (Okello *et al.*, 2015); *T. b. brucei* and *T. b. rhodesiense* cause only mild illness in indigenous breeds with infection often going undetected (Fyfe *et al.*, 2017). It is estimated that 20% production

losses are experienced due to mortality and the reduction of calving rate, draft power and the production of meat and milk due to AAT infections (Swallow, 2000).

1.2 Trypanosome classification

Trypanosomes are protozoan parasites and belong to the order Kinetoplastida and genus Trypanosoma. The genus comprise the Stecoraria and Salivaria groups (Haag *et al.*, 1998). Trypanosomes belonging to the Stecoraria group, migrate to the hind gut of the vector and are transmitted through faeces to the mammalian host, whilst those in the Salivaria group, migrate from the gut to the mouth parts of the vector and are transmitted through the vector's saliva (Lucius *et al.*, 2017). The salivarian parasite, *T. equiperdum*, is an exception as it is transmitted during copulation (Brun *et al.*, 1998). Nine trypanosomal clades have been identified based on the heat shock protein (HSP) 60 molecular marker (Fraga *et al.*, 2016); namely, *T. cruzi, T. rangeli*, *T. lewisi, T. theileri, T. grayi, T. carassi, Trypanozoon T. congolense* and *T. vivax* as outlined in Table 1.1. The *Trypanozoon* clade includes *T. b. gambiense, T. b. rhodesiense, T. b. brucei, T. evansi and T. equiperdum.*

Table 1.1: Taxonomy of the trypanosomatidae

Group	Subgenus	Species	Subspecies	Clade[a]
Stecoria	Schizotrypanum[#]	*T. cruzi*		*T. cruzi*
	Herpetosoma	*T. rangeli*		*T. rangeli*
		T. lewisi		*T. lewisi*
		T. musculi		
	Megatrypanum[#]	*T. theileri*		*T. theileri*
		T. grayi		*T. grayi*
		T. carassi		*T. carassi*
		T. avium		Avian[b]
Salivaria	Trypanozoon	*T. brucei*	*T. b. gambiense* *T. b. rhodesiense*	*Trypanozoon*
		T. evansi		
		T. equiperdum		
	Nannomas	*T. congolense*		*T. congolense*
		T. simiae		
		T. godfreyi		
	Dutonella	*T. vivax*		*T. vivax*
		T. uniforme		
	Pycnomonas	*T. suis*		

Adapted from Baral (2010), Fraga *et al.* (2016)[a] and Simpson *et al*. (2006)[b].
[#]Only a few species belonging to the Schizotrypanum and Megatrypanum subgenus are given.

1.3 Parasite structure and genomic organisation

The elongated shape of the trypanosomal trypomastigotes and epimastigotes is defined by a polarised microtubule cytoskeleton, which remains intact throughout the cell cycle (Matthews, 2005). The trypanosomal parasites possess a number of single copy organelles which are precisely positioned within the cytoskeleton; namely, the nucleus, flagellar pocket, basal body, kinetoplast, mitochondrion, nucleus, Golgi apparatus and the flagellum (Matthews, 2005), which is attached to the body of trypomastigotes and epimastigotes (Fig. 1.2).

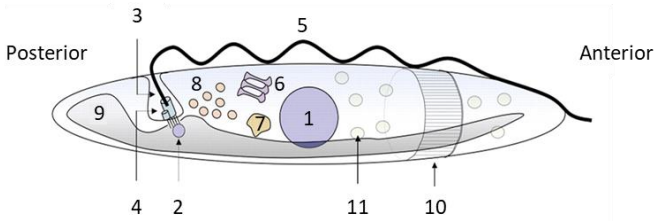


Figure 1.2: The cellular structure of *T. b. brucei*. The cellular structures found within the *T. b. brucei* trypomastigote are shown and numbered as follows: 1, nucleus; 2, kinetoplast; 3, flagellar pocket; 4, basal body and probasal body; 5, axoneme and paraflagellar rod; 6, Golgi apparatus; 7, lysosome; 8, endosomes; 9, mitochondrion; 10, microtubule cytoskeleton and 11, glycosomes (Matthews, 2005).

The single kinetoplast comprises the mitochondrial genome and consists of approximately 50 copies of maxicircles and 10,000 copies of minicircles (Klingbeil and Englund, 2004). Mitochondrial proteins are encoded by the maxicircles, whilst short guide RNAs are encoded by the minicircles (Klingbeil and Englund, 2004). The maxicircle transcripts are post-transcriptionally edited by the incorporation or deletion of uridines using the minicircles as templates (Klingbeil and Englund, 2004). The replication of the kinetoplast is coordinated with that of the nucleus (Matthews, 2005). The nuclear genome of *T. brucei* contains 11 megabase chromosomes and more than 100 minichromosomes which are approximately 50 kb in size (Wickstead *et al.*, 2004; Cross *et al.*, 2014). The minichromosomes function to provide a large repertoire of telomeric variable surface glycoproteins, VSGs (Pays, 2006; Hovel-Miner *et al.*, 2016). The genes are arranged in polycistronic arrays, which are not organised into operons, but are rather differentially expressed through the life cycle (Matthews, 2005).

1.4 VSG and antigenic variation

A dense ~15 nm-thick coat of VSGs is present on the surface of the parasite and protects against complement-mediated lysis and hides invariant proteins from

recognition by the host's immune system (Engstler *et al.*, 2004; Field and Carrington, 2009; Schwede *et al.*, 2011). Bloodstream trypomastigote parasites possess about 10^7 copies of the VSG, which constitutes approximately 5% of the total cellular protein (Turner, 1985).

The VSG molecule is composed of a large, alpha-helical N-terminal domain which interacts with the host immune system (Blum *et al.*, 1993), and a small C-terminal domain which links the VSG dimer to the plasma membrane via a glycosylphosphatidylinositol (GPI) anchor (Jones *et al.*, 2008). The VSG repertoire of *T. b. brucei* bloodstream trypomastigotes, comprising more than 1 000 VSGs (Berriman *et al.*, 2005), is significantly larger than the 12 VSGs found in procyclic trypomastigotes (Lenardo *et al.*, 1984).

As VSGs are highly immunogenic, an immune response is generated by the host to eliminate the parasite. However, the periodic switching of the expressed VSGs, with distinct antigen variants, allows for immune evasion, resulting in an undulating parasitaemia profile, characteristic of trypanosome infections (MacGregor *et al.*, 2012). The process of antigenic variation is the result of either transcriptional switching between VSG expression sites (ESs), or homologous recombination between the active VSG gene and another VSG gene from the repertoire (Pays *et al.*, 2006). The genome of *T. b. brucei* encodes as many as 1 700 VSGs, most of which are pseudogenes which require assembly into functional mosaics (Marcello and Barry, 2007). This implies that the VSG repertoire is essentially infinite (Taylor and Rudenko, 2006).

A single VSG is expressed at any one given time as only a single VSG allele and can be transcribed from the VSG repertoire (Pays *et al.*, 2006). The transcription of VSG occurs at one of several VSG ESs (Pays *et al.*, 2006). Each VSG ES has polycistronic transcription units which contain expression-site-associated genes (*ESAG*s) together with the VSG gene (Pays *et al.*, 2001). The *T. b. brucei* genome codes for a set of 15 to 20 similar, but not identical ESs (Pays *et al.*, 2006).

The *serum resistance associated* (*SRA*) gene encodes a truncated VSG where the region encoding the surface-exposed epitopes is missing (Vanhamme *et al.*, 2001). The R-ES is the ES which contains the *SRA* gene (De Greef *et al.*, 1989) and is not active when the parasite is present in the serum of non-human vertebrates (Vanhamme *et al.*, 2000). Only when *T. b. rhodesiense* is exposed to human serum, is the R-ES selected and the transcription of *SRA* triggered (Pays *et al.*, 2006). Animal

infective trypanosome parasites do not possess the *SRA* gene and are unable to infect humans as they are lysed in human serum.

1.5 Life cycle of the trypanosomal parasite

The complex life cycle of the trypanosome requires various adaptations to survive in the mammalian bloodstream and in the different compartments within the tsetse fly (Matthews, 2005). Stage-specific changes to basic cell biological processes are required for differentiation (Matthews, 2005). The life cycle of *T. b. brucei* is detailed in Fig. 1.3.

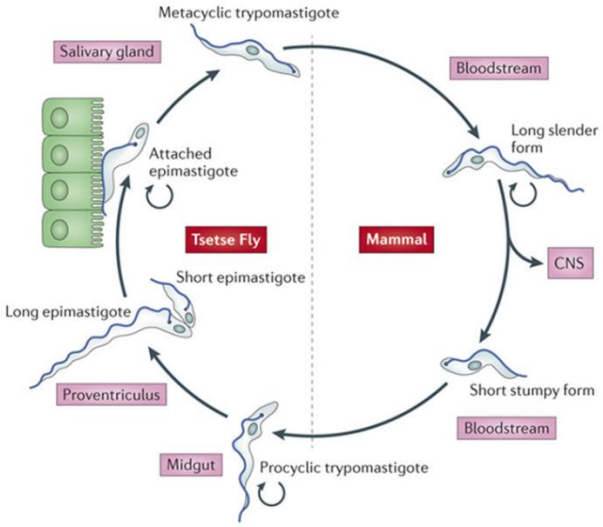


Figure 1.3: Life cycle of *T. b. brucei* in the mammalian host and insect vector. The circular arrows indicate that the form is proliferative. CNS, central nervous system (Langousis and Hill, 2014).

The mammalian host is infected through the bite of a tsetse fly which harbours growth-arrested *T. b. brucei* metacyclic trypomastigotes. Thereafter, the metacyclic trypomastigotes differentiate into proliferative slender forms which are able to establish and maintain infection in the host circulation as bloodstream trypomastigotes (Matthews, 2005) through evasion of the host antibody responses using antigenic variation (Pays *et al.*, 2004). Once in the bloodstream, the long slender forms differentiate into short stumpy forms, which are cell-cycle arrested and pre-adapted for survival in the tsetse fly, facilitated by a quorum sensing-like mechanism (MacGregor *et al.*, 2011). This transition between slender and stumpy forms, together with the process of antigenic variation, prolongs the infection and increases the potential for further parasite transmission (MacGregor *et al.*, 2012). In the second stage of *T. brucei*

infection, parasites penetrate the blood brain barrier and enter the central nervous system (Langousis and Hill, 2014).

A blood meal, from an infected mammalian host, allows for the parasites to establish in the tsetse midgut, through the differentiation of short stumpy forms into proliferative procyclic trypomastigotes. The differentiation of bloodstream trypomastigotes to procyclic trypomastigotes is characterised by morphological changes, switching from glycolysis to oxidative phosphorylation, and the replacement of the VSGs with procyclins (Maudlin and Welburn, 1994), which are also GPI anchored (Roditi and Liniger, 2002). Procyclic trypomastigotes migrate from the midgut to the salivary gland, and *en route*, undergo extensive restructuring and asymmetric division to generate both long and short epimastigotes which are covered in brucei alanine-rich protein (BARP) (Urwyler *et al.*, 2007). The short epimastigotes attach to the epithelial cells in the salivary gland and are able to replicate via asymmetric division to generate metacyclic trypomastigotes (Matthews, 2005). The metacyclic trypomastigotes regain a coat of VSGs and are released into the salivary gland lumen where they can be transferred to the mammalian host; thereby, completing the life cycle (Rotureau *et al.*, 2012).

Similar, but non-identical *in vivo* life cycles are shared between *T. congolense* and the *T. brucei* subspecies. The bloodstream trypomastigotes of *T. brucei* occur as proliferative long slender forms as well as a non-proliferative short stumpy forms, whereas those from *T. congolense* have the corresponding proliferative forms but the stumpy forms are less obvious (Vickerman, 1985). The epimastigotes of *T. congolense* are directed to the tsetse fly's proboscis and mouth parts, but not to the salivary glands as for *T. brucei* epimastigotes, where they differentiate into non-proliferative, infective metacyclic trypomastigotes (Roditi and Lehane, 2008). In the case of *T. vivax*, the procyclic trypomastigote form is skipped as the parasite does not enter the insect gut (Rotureau and Van Den Abbeele, 2013). Instead, the bloodstream trypomastigotes differentiate directly into epimastigotes within the tsetse proboscis (Rotureau and Van Den Abbeele, 2013).

1.6 Immune response

African trypanosomes are extracellular parasites, and as such, encounter both the innate and the adaptive immune responses of the host. The VSG molecules provide a defence barrier against both innate and specific immune effectors (Vickerman and

Luckins, 1969). Polyclonal B-cell activation is a characteristic of both bovine and murine trypanosomosis, whereby increased numbers of B-cells and plasma immunoglobulin levels is evident (Luckins and Mehlitz, 1976).

Proliferation of B-cells, producing a lytic antibody response, is caused by the immunogenic VSGs as illustrated in Fig. 1.4 (Pays *et al.*, 2001). During the early stage of infection, trypanosome-specific antibodies can mediate clearance of parasites and are able to neutralise parasite products (Dempsey and Mansfield, 1983). However, a significant proportion of these antibodies are either polyspecific or auto-reactive (Kobayakawa *et al.*, 1979). The suppression of B cell memory results in the total absence of IgG responses and a strongly reduced IgM response in the later stages of infection (Hudson and Terry, 1979; Radwanska *et al.*, 2008).

The VSG-antibody complexes facilitate opsonisation and lysis via the complement system of the parasites expressing the VSG against which the response was triggered. Thus, by the process of antigenic variation, trypanosomes avoid detection by host antibodies (Horn and McCulloch, 2010). Through the reduction of parasite numbers, prolonged infection is ensured.

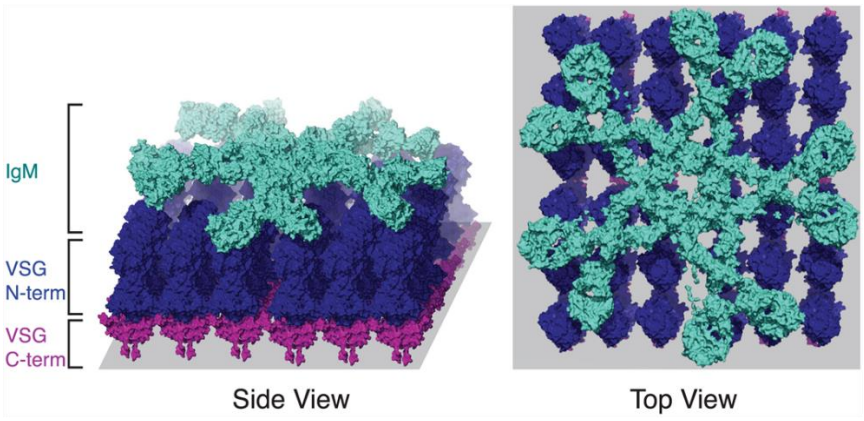


Figure 1.4: Lytic antibody response to the immunogenic VSGs. The hypothetical interaction between the IgM (turquoise) and the N-terminus of the immunogenic VSG (dark blue) molecules (Mugnier *et al.*, 2016).

The initial immune response to trypanosome infection is the initiation of an acute inflammatory response via the innate system, which affects B- and T-cell responses to parasite antigens, whereby cytokines and chemokines are secreted to prompt clearance of the infecting parasite (Mansfield and Paulnock, 2005). Natural killer cells may also participate in the initiation of the inflammatory response through the synthesis of cytokines and chemokines such as interferon (IFN)-ɣ and tumor necrosis factor (TNF)-α (Vincendeau and Bouteille, 2006). However, the parasite can interfere with

the cytokine network, using the cytokines as growth factors, and modifying the effector functions of the immune system (Hide *et al.*, 1989).

The first response of the host immune system consists of classically activated macrophages secreting pro-inflammatory molecules such as TNF, IL-1, IL-6 and nitric oxide (NO) (Duxbury *et al.*, 1972; Mosser and Roberts, 1982; Pan *et al.*, 2006). The classically activated macrophages phagocytose antibody-opsonised parasites (Shi *et al.*, 2004) as well as secrete trypanotoxic molecules such as TNF and NO (Mabbott *et al.*, 1994; Sternberg and Mabbott, 1996; Kaushik *et al.*, 2000) which are involved in the control of the first peak of parasitaemia.

Despite the benefit of the initial inflammatory response, sustained inflammation can cause further damage to the mammalian host. Thus, a reduction in inflammation is required by the down regulation of the classically activated macrophages and their pro-inflammatory cytokines. The production of type II cytokines such as IL-4, IL-10 and IL-13 are able to modulate the macrophages to become alternatively activated macrophages, which are anti-inflammatory (Baral, 2010).

1.7 Control, diagnosis and chemotherapy for trypanosomosis

Despite an increase in health facilities throughout endemic countries that offer diagnosis and treatment for HAT, access to these facilities is limited due to the remote locations of the areas of incidence, which are often also in war zones (Franco *et al.*, 2017). In order to maintain and further reduce the numbers of HAT and AAT cases, innovative and cost-effective strategies are required. This includes the development of new tools for diagnosis, vector and reservoir control as well as effective trypanocides (Franco *et al.*, 2017).

1.7.1 Control

The control of AT can be accomplished by reduction of parasite transmission through the control of the insect vector, as well as case detection and treatment (Büscher *et al.*, 2017). Due to the existence of a wildlife reservoir, the elimination of HAT is unlikely, if the insect vector is not removed. However, the removal of tsetse populations has been achieved in less than 2% of the 10 million km^2 estimated to be infested (Bouyer *et al.*, 2013), and AAT integrated management has rarely been sustained (Vreysen *et al.*, 2013; Shaw *et al.*, 2015).

A number of trypanosomosis and tsetse control methods have been developed and implemented with varied levels of success in a number of endemic regions (Muhanguzi *et al.*, 2015). Methods for the control of the insect vector include both stationary and mobile baits in the form of fly traps (Vale and Torr, 2004) and insecticide-treated cattle (Allsopp and Hursey, 2004), aerial spraying (Kgori *et al.*, 2006), sterile insect technique (Vreysen *et al.*, 2000) and the administration of trypanocides to cattle (Holmes *et al.*, 2004).

The 86% reduction in *T. b. gambiense* infections, between 2000 and 2014, can be attributed to consistent active and passive surveillance, screening approximately 2 million people per year (Franco *et al.*, 2017). Vector control has been initiated in some *T. b. gambiense* endemic regions and is believed to have contributed to the reduction of reported cases in countries such as Guinea (Courtin *et al.*, 2015) and Chad (Franco *et al.*, 2017). Infections with *T. b. rhodesiense* represent an average of 2.7% of the total number of reported HAT cases between 2000 and 2014 (Franco *et al.*, 2017). The wildlife reservoir, which is difficult to manage, is the cause of the scattered but constant *T. b. rhodesiense* cases reported (Simarro *et al.*, 2010). Due to the faster disease progression and poor effectiveness of active screening of *T. b. rhodesiense* infections, the rate of under-detection is, in all likelihood, higher than that of *T. b. gambiense* infections (Franco *et al.*, 2017).

The most widely used AAT control method is trypanocides (Grant, 2001; Holmes *et al.*, 2004; Torr *et al.*, 2007), as livestock keepers in the tsetse infested areas of Africa, consider this to be a rapid and inexpensive means of control compared to that of tsetse control methods (Torr *et al.*, 2007). This method requires multiple trypanocide administrations as the currently used trypanocides afford only a short period of protection (Bourn *et al.*, 2005). However, there is a high risk of the development of drug resistance through the extensive use of trypanocides (Geerts *et al.*, 2001; Holmes *et al.*, 2004).

In an attempt to control *T*. *b*. *rhodesiense* infections in livestock, mass administration of isometamidium chloride followed up by routine application of insecticides to protect cattle from tsetse flies (Welburn *et al.*, 2006a; Kajunguri *et al.*, 2014), has been implemented across large endemic focus areas in Uganda (Fyfe *et al.*, 2017). This intervention allowed for the protection of the human population from HAT infections as well as the improvement of the health and productivity of livestock (Okello *et al.*, 2015).

1.7.2 Diagnosis

Ideal diagnostics in resource-limited settings should meet the WHO quality-ASSURED criteria (Peeling *et al.*, 2006). This includes affordability, sensitivity, specificity, user friendly with minimal training, rapid and robust, equipment free and deliverable to end users.

Diagnosis of trypanosomal infections require serological screening, followed by microscopic confirmation, and finally in the case of HAT, disease staging (Chappuis *et al.*, 2005). Diagnosis of *T. b. gambiense* infections can be performed in 882 fixed health facilities, of which, 639 can perform serological testing, 326 are equipped to perform parasitological diagnosis and only 262 can perform disease staging (Franco *et al.*, 2017). Diagnosis of *T. b. gambiense* infections in the six endemic countries, i.e. Kenya, Malawi, Uganda, United Republic of Tanzania, Zambia and Zimbabwe, are offered by 111 health facilities, all of which are able to perform clinical diagnosis (Franco *et al.*, 2017). Parasitological diagnosis is only offered by 49 and disease staging by 34 of the 111 facilities (Franco *et al.*, 2017).

As the direct result of the highly immunogenic VSG molecules, high concentrations of VSG-specific IgG and IgM antibodies are produced (Crowe *et al.*, 1984). Serological screening using the card agglutination test, CATT, detecting host antibodies against the LiTat 1.3 variant VSG of *T. b. gambiense* (Magnus *et al.*, 1978) is the first step in diagnosis (Chappuis *et al.*, 2005) and characterised by a 69-100% sensitivity and 84-99% specificity (Mitashi *et al.*, 2012). The immune trypanolysis test is based on complement-mediated lysis of trypanosomes (Chappuis *et al.*, 2005). Despite being the gold standard in HAT and AAT diagnostics (Van Meirvenne *et al.*, 1995) with a 97.2-100% sensitivity and 100% specificity, the test is expensive and associated with a high infection risk to personnel (OIE, 2013).

Microscopic examination of lymph node aspirates, blood or cerebrospinal fluid (CSF) from sero-positive patients provides direct evidence for trypanosome infection (Chappuis *et al.*, 2005). Parasite detection can be labour intensive and failure to detect parasites does not rule out infection. With the aid of centrifugation techniques to concentrate parasites in the buffy coat fraction, prior to microscopic analysis, sensitivity can be improved (Woo, 1970). In the absence of a centrifuge, wet and thin blood films are the methods of choice for the confirmation of AT infection. However, parasitological tools are often not sensitive enough to detect parasites which may be present in bodily

fluids in low numbers (Field *et al.*, 2009), and is estimated to fail to detect between 20 and 30% of trypanosome infections (Robays *et al.*, 2004).

Following parasite visualisation, HAT disease staging is performed by the examination of CSF samples. The presence of parasites in CSF is indicative of second stage infection (Chappuis *et al.*, 2005). However, the number of circulating parasites in the CSF may be very low, which may cause false negatives. Thus, sensitive diagnostic tools are required for the detection of low levels of parasites at various points during the infection period.

Population screening using rapid diagnostic tests (RDTs) are able to detect antibodies in human finger-prick blood samples (Posthuma-Trumpie *et al.*, 2009). These tests meet the ASSURED criteria as they are cost effective, simple to use and stable for distribution into remote areas of disease incidence. The SD BIOLINE HAT (Standard Diagnostics, Fig. 1.5, panel A) and the Sero-K-SeT (Coris Bioconcept) have been developed for the diagnosis of *T. b. gambiense* infections based on the detection of anti-trypanosome antibodies against the native *T. b. gambiense* LiTat 1.3 and LiTat 1.5 antigens (Büscher *et al.*, 2013). However, no RDT for *T. b. rhodesiense* has been developed (Büscher *et al.*, 2017). Comparison of the diagnostic capabilities of the BIOLINE and Sero-K together with the CATT and trypanolysis methods, using sera from endemic regions: Angola, Democratic Republic of the Congo, and Central African Republic, have been reported by a number of research groups for *T. b. gambiense* infections (Büscher *et al.*, 2013; Büscher *et al.*, 2014; Jamonneau *et al.*, 2015; Bisser *et al.*, 2016; Lumbala *et al.*, 2017). The general consensus is that these RDTs are as sensitive and specific as the CATT and trypanolysis methods, and can, thus, be used for active screening in remote areas more easily than the CATT which requires a cold chain and electricity.

As with the immune trypanolysis test, the use of native antigens in BIOLINE and Sero-K RDTs poses a risk to personnel due to the culture of human infective trypanosomes. As such, the application of recombinant antigens in an RDT format is being investigated. A study using dual-antigen lateral flow test prototypes with sera from the Democratic Republic of the Congo and WHO HAT Specimen Biobank (as a virtual field trial) showed promising results (Sullivan *et al.*, 2014). Several prototypes have been developed using recombinant as well as native antigens. The prototype developed by Sullivan *et al*. (2014) was characterised by a sensitivity of 97% and specificity of 83% when using recombinant soluble VSG117/MiTat 1.4 and invariant surface glycoprotein

(ISG)-65. A sensitivity 82% and specificity 95% was obtained with the prototype reported by Sternberg *et al*. (2014) using recombinant LiTat 1.3 and -1.5, and an 88% sensitivity and 94% specificity for that developed by BBI Solutions (Cardiff, Wales), using recombinant ISG65 and native MiTat 1.4 (Sternberg *et al.*, 2014). These prototypes have comparable sensitivity and specificity to that of the BIOLINE (89-99% and 88-98%) and CATT (69-95% and 95-98%) as measured in endemic areas (Büscher *et al.*, 2013; Büscher *et al.*, 2014; Bisser *et al.*, 2016; Lumbala *et al.*, 2017). The application of LiTat 1.3 and -1.5, recombinantly expressed in *L. tarentolae*, in a diagnostic ELISA demonstrated that these antigens had a comparable sensitivity and specificity to that of the native LiTat 1.3 and -1.5 (Rooney *et al.*, 2015). These results highlight the applicability of recombinantly expressed antigens in the generation of accurate HAT RDTs.

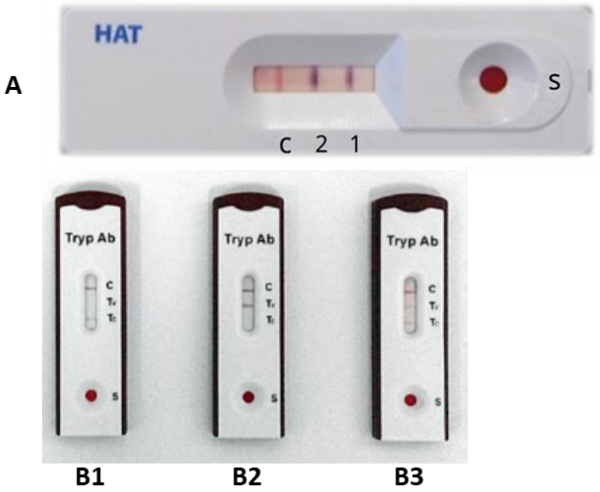


Figure 1.5: Human and animal RDTs for the diagnosis of African trypanosomosis. (A) Positive test for *T. b. gambiense* using the SD BIOLINE HAT RDT where S, sample well; 1, LiTat1.3; 2, LiTat1.5 and C, control (Bisser *et al.*, 2016). Positive AAT prototype for **(B1)** *T. congolense* infection, **(B2)** *T. vivax* infection and **(B3)** mixed *T. congolense* and *T. vivax* infection, where C, control; Tv, *Tvi*GM6 antigen and Tc, *Tco*CATB antigen (Boulangé *et al.*, 2017).

Serological diagnostic tools, such as the CATT, are limited to detection only and are not suitable for follow-up evaluations due to the persistence of anti-trypanosome antibodies after cure (Lejon *et al.*, 2010). To distinguish between active and cured AT, a serological diagnostic test would have to be able to detect the antigens produced by the parasite rather than the IgM antibodies produced against the parasite (Tiberti *et al.*, 2013). Trypanosomal antigens, which are circulating during active infections, are, thus, potential targets for diagnostics and even chemotherapy. Pathogeno-proteomics is considered a promising method for the identification of new diagnostic markers and new therapeutic targets (Hölzmuller *et al.*, 2008). Using this method, diagnostic

antigens in *T. b. gambiense* (Sullivan *et al.*, 2013) and *T. vivax* have been identified for use in diagnostics (Fleming *et al.*, 2016).

The GM6 antigen has been identified as a suitable diagnostic antigen for AAT diagnosis. The tandem repeat antigen is conserved amongst the salivarian trypanosomal parasites and the recombinantly produced antigen was applied in an indirect ELISA for *T. vivax* diagnosis (Pillay *et al.*, 2013) and an RDT format for *T. evansi* diagnosis (Nguyen *et al.*, 2015). However, this *T. evansi* RDT has not been commercialised. The use of the *Tco*CATB (Mendoza-Palomares *et al.*, 2008) and *Tvi*GM6 (Pillay *et al.*, 2013) antigens in a dual-antigen RDT prototype, was characterised by a sensitivity of 96% and a 92% specificity for the detection of both *T. congolense* and *T. vivax* infections (Fig 1.5, panel B) (Boulangé *et al.*, 2017).

Two related members of the *T. vivax* ISG family were shown to have good diagnostic applicability and the RDT prototype was characterised with a 92% sensitivity and 89.9% specificity using sera from experimental infections (Fleming *et al.*, 2016). Taken together with the success of the ISG65 in the HAT RDT prototypes (Sternberg *et al.*, 2014; Sullivan *et al.*, 2014), these ISGs have great diagnostic potential for animal infections.

Diagnostic methods, which do not meet the ASSURED criteria, include PCR and various modifications thereof. Despite being highly sensitive (70-100%) and specific (91.4-100%), DNA-based methods are not commonly used for active case finding in rural HAT endemic regions, due to the expense and requirement for sophisticated equipment and trained personnel (Mitashi *et al.*, 2012). Instead, DNA-based methods are useful for population screening, with amplification of the 177 bp satellite DNA (Wickstead *et al.*, 2004) and *SRA* gene (Welburn *et al.*, 2001b) for HAT, 18S ribosomal sub-unit (Geysen *et al.*, 2003) and the internal transcribed spacer (ITS1) of ribosomal DNA (Desquesnes *et al.*, 2001) for AAT. The ITS1 PCR method is able to successfully distinguish between *T. brucei* spp., *T. congolense, T. simiae, T. godfreyi* and *T. vivax* infections in cattle and camels (Njiru *et al.*, 2005). The finding that the ITS PCR method is suitable for large-scale epidemiological studies (Thumbi *et al.*, 2008), and was used for the evaluation of trypanosome prevalence in cattle (Ahmed *et al.*, 2013) and in tsetse flies (Isaac *et al.*, 2016) in endemic areas. This data can serve as an indicator of where parasite control needs to be implemented. This was the case where the screening of goats and cattle in Zambia indicated that prevention and control strategies had reduced the number of infections (Musinguzi *et al.*, 2016).

1.7.3 Chemotherapies

The available chemotherapies used for the treatment of HAT are highly toxic, require trained healthcare providers and drug resistance has been reported (Garcia-Salcedo *et al.*, 2014). The ideal trypanocide should be affordable, safe, effective and require simple administration (Kuzoe, 1993). In addition, the identity of the infecting parasite and disease stage needs to be considered prior to treatment. The currently used disease confirmation and staging are complex and painful to the patient, which results in under-reporting of cases (Odiit *et al.*, 2005; Acup *et al.*, 2017).

First and second stage *T. b. gambiense* infections are treated with pentamidine and *e*flornithine, respectively (Kennedy, 2013). More recently, the use of nifurtimox-eflornithine combination therapy (NECT) was recommended due to increased efficacy, decreased dosing and a reduction in emergence of drug resistance, despite the higher cost (Priotto *et al.*, 2009; Lutje *et al.*, 2010). Suramin and the organo-arsenical, melarsoprol, are administered for first and second stage *T. b. rhodesiense* infections, respectively (Kennedy, 2013). Each trypanocide causes adverse side effects, the worst of which is caused by melarsoprol, with 5 to 10% of treated patients succumbing to reactive encephalopathy (Burri, 2010).

Target-based and phenotypic screening have been employed to identify both attractive molecular targets and cell permeable inhibitors which are toxic to the parasite but not to the mammalian host (Matthews, 2015). The identification and development of the orally effective drugs fexinidazole and benzoxaborole SCYX-7158 have been the direct result of phenotypic screening (Maser *et al.*, 2012). Fexinidazole has almost completed phase 2/3 clinical trials (Tarral *et al.*, 2014), whilst SCYX-7158 has just entered phase 2/3 trials (Jacobs *et al.*, 2011).

Due to the decrease in the number of reported HAT cases (Franco *et al.*, 2017), the human impact of AT is largely indirect, through the effects on livestock (Matthews, 2015). However, research into the biochemistry and drug susceptibility of the causative agents of AAT is limited, and their importance is often overlooked.

The currently used chemotherapies for AAT are diminazene aceturate (Berenil®), isometamidium chloride (Samorin®) and homidium chloride (Novidium®) (Geerts *et al.*, 2001). The indiscriminate use of trypanocides have resulted in reduced effectiveness and the emergence of drug resistance (Van den Bossche *et al.*, 2000; Geerts *et al.*, 2001; Delespaux *et al.*, 2002). Since the introduction of isometamidium chloride in 1961 (Berg *et al.*, 1961), no new AAT treatment has been introduced (Giordani *et al.*,

2016). As such, it is not surprising that drug resistance is on the increase (Van den Bossche and Delespaux, 2011) and has been reported in 21 African countries (Delespaux and de Koning, 2007; Chitanga *et al.*, 2011). A strategy to prevent the development of drug resistance is the use of isometamidium chloride or homidium chloride, together with diminazene aceturate as a sanative pair (Geerts *et al.*, 2001; Holmes *et al.*, 2004; Chitanga *et al.*, 2011).

It has been suggested that to prevent reinfection, trypanocides together with veterinary insecticides are required (Kajunguri *et al.*, 2014; Muhanguzi *et al.*, 2014), which is an economical option for HAT control as well as the control of tick-borne diseases and AAT (Muhanguzi *et al.*, 2014).

1.8 Peptidase virulence factors

Due to the ability of trypanosomes to undergo antigenic variation and elimination of host B-cell memory (Radwanska *et al.*, 2008), the development of a vaccine is highly unlikely (La Greca and Magez, 2011). Thus, prevention, through vector control, accurate diagnosis and timely treatment are the only options for the control of AT. In the absence of a vaccine, an anti-disease strategy was proposed for the neutralisation of the pathogenic effects of the trypanosomal parasite (Antoine-Moussiaux *et al.*, 2009).

Pathogeno-proteomics entails an analysis of the interactions between the parasite, mammalian host and insect vector, towards characterisation of the mechanisms by which the disease progresses (Hölzmuller *et al.*, 2008). This analysis identifies molecular targets which are essential for pathogenesis, and are termed virulence factors (Antoine-Moussiaux *et al.*, 2009). Parasite peptidases are thought to play direct roles in disease pathogenesis as they are involved in host invasion, migration, metabolism and immune evasion (McKerrow *et al.*, 2006). The cathepsin-L-like peptidases (Steverding *et al.*, 2012), metacaspases (Helms *et al.*, 2006), oligopeptidases (Morty *et al.*, 2001) and the aurora kinases (Tu *et al.*, 2006) are some of the trypanosome virulence factors identified to date.

In addition to being potential diagnostic antigens, the virulence factors are attractive targets for the development of novel chemotherapies (Drag and Salvesen, 2010). Recently, the kinetoplastid proteasome was genetically and chemically validated as a promising target for the development of novel chemotherapies for the treatment of trypanosomosis, Chagas disease and leishmaniasis (Khare *et al.*, 2016).

1.8.1 Metacaspases

Caspases (cysteine-dependent aspartate specific proteases) are found only in the metazoan kingdom, whilst the other kingdoms possess the structurally related metacaspases, MCAs (Uren *et al.*, 2000). Due to their structural similarity to the caspases, specifically the catalytic dyad and classic caspase-haemoglobinase fold (Aravind and Koonin, 2002) the MCAs have been assigned to clan CD, and the C14B subfamily of cysteine peptidases (Rawlings *et al.*, 2016). The MCAs differ from caspases in their preference for basic Arg and Lys residues over acidic Asp residues at the P_1 substrate position, together with a low shared sequence homology and their inability to form dimers (McLuskey *et al.*, 2012).

A highly selective S_1 pocket ensures the strict substrate specificity of clan CD peptidases. This feature permits their involvement in specific processing events, with little chance of random peptidolytic degradation. An example of this specificity is evident in the signalling pathways involving caspases during their participation in apoptosis and inflammation (Pop and Salvesen, 2009).

1.8.1.1 Structure of metacaspases

The MCAs are divided into type I, to which the ancestral form found in unicellular organisms belongs, and type II, which have evolved for larger, more complex organisms (Lam and Zhang, 2012). More recently, type III metacaspases, found only in algae which have undergone secondary endosymbiosis, have been discovered (Choi and Berges, 2013). Type I MCAs are typically found in plants, yeasts and protozoa whilst type II MCAs are found only in plants. The most studied MCAs are from *Arabidopsis thaliana*, which possess both type I MCAs 1 to 3 and type II MCAs 4 to 9.

Similar to the caspases, the MCAs possess a large catalytic domain (p20), within which the catalytic dyad is found, and a smaller N-terminal domain (p10) (Fig. 1.6). The type II MCA is an exception, where the p10 domain is found at the C-terminus. Unlike caspases, type I MCAs possess a short linker region and type II MCAs a long linker region between the p20 and p10 domains. An N-terminal domain, which is present in initiator caspases, is found in type I MCAs, which in plants contains a zinc finger motif (Vercammen *et al.*, 2004; Tsiatsiani *et al.*, 2011). Conversely, type II MCAs lack the N-terminal domain and are similar to execution caspases, by way of their dependency on autoprocessing for activation (Vercammen *et al.*, 2004; Tsiatsiani *et al.*, 2011).

The MCAs are a functionally diverse family of peptidases, and may harbour a capacity for multifunctionality within organisms which possess only a single MCA copy. The single copy MCA from *Leishmania* spp. and the trypanosomal MCA5s possess a Pro-, Gln- and Tyr-rich C-terminal extension, unlike the short C-terminal extension of the multicopy genes, *Tbb*MCA 1 to 4 (Mottram *et al.*, 2003) and the 16 copies of *Tcr*MCA3 (Kosec *et al.*, 2006). The C-terminal domain of *Lmj*MCA interacts with proteins involved in stress regulation (Casanova *et al.*, 2015). This demonstrates that the Pro-rich C-terminal domain of the MCAs could be responsible for protein-protein interactions as suggested by Kay *et al*. (2000).

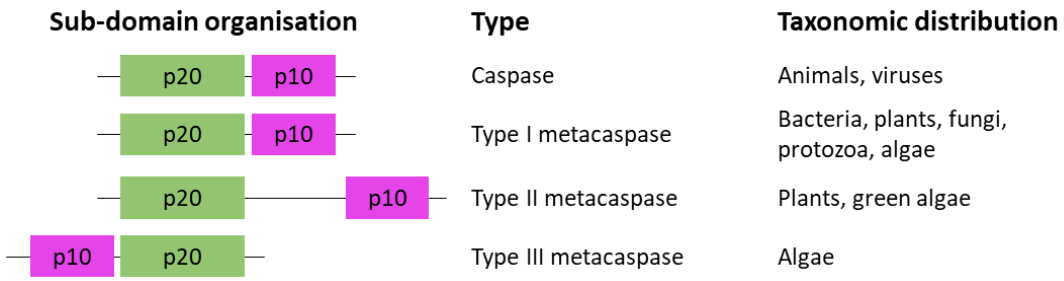


Figure 1.6: Domain organisation of the caspases and metacaspases. The catalytic domain (p20), in which the catalytic dyad is found, is highlighted in green, whereas the smaller, p10 domain, is highlighted in pink. The domains are not drawn to scale and the additional domains are not shown in this instance.

1.8.1.2 Functions of metacaspases

The study by Madeo (2002) identified the *Saccharomyces cerevisiae* MCA (YCA1) as a positive regulator of programmed cell death (PCD). YCA1 has been implicated in both cell cycle regulation (Lee *et al.*, 2008) and the clearance of insoluble protein aggregates (Lee *et al.*, 2010). It has been demonstrated that the single copy *Lmj*MCA has a role in PCD induced by oxidative stress (Zalila *et al.*, 2011) and is also required for cell cycle progression (Ambit *et al.*, 2008). In addition, an antagonistic relationship exists between two of the *A. thaliana* MCAs in the hypersensitive response cell death pathway, whereby *At*MCA1 promotes cell death and *At*MCA2 functions as a negative regulator. Together, this highlights the potential of the action of MCAs in key regulatory processes (Coll *et al.*, 2010).

An RNAi study revealed that the pseudopeptidase, *Tbb*MCA4, functioned in both cell cycle progression and parasite virulence during mammalian infection (Proto *et al.*, 2011). It was demonstrated that *Tbb*MCA4 was processed by *Tbb*MCA3, both of which

are palmitoylated (Emmer *et al.*, 2009), and may comprise a MCA peptidolytic cascade. (Proto *et al.*, 2011).

Irrespective of the elucidation of the role of MCA, the lack of a mammalian homologue and the MCAs' involvement in the parasite cell cycle and in cell death pathways, make the MCAs attractive targets for the development of novel chemotherapies for trypanosomosis, Chagas disease and leishmaniasis (Ambit *et al.*, 2008; McKerrow *et al.*, 2008; Alvarez *et al.*, 2011). In addition, the presence of MCAs in the plant and fungi kingdoms, make them attractive targets for the development of pesticides and fungicides (Lam and Zhang, 2012).

1.8.2. Oligopeptidase B

The serine peptidase, oligopeptidase B (OPB), is present in the cytosol of trypanosomes and bacteria (Burleigh *et al.*, 1997; Morty *et al.*, 2005a; Rea and Fülöp, 2006), and has been identified as an important virulence factor in AAT caused by *T. b. brucei, T. congolense* and *T. evansi* (Troeberg *et al.*, 1996; Morty *et al.*, 1999; Morty *et al.*, 2001; Morty *et al.*, 2005a). The trypanocidal drugs suramin (Morty *et al.*, 1998), pentamidine (Ashall, 1990; Morty *et al.*, 1998) and diminazene (Morty *et al.*, 1998), were shown to inhibit *Tbb*OPB. In addition, irreversible inhibitors of *Tbb*OPB exhibited anti-trypanosomal activity *in vitro* and *in vivo* (Morty *et al.*, 2000), and OPB can therefore be regarded as a potential target for the development of novel trypanocides (Coetzer *et al.*, 2008; Canning *et al.*, 2013). However, despite a 60% protein sequence similarity between the OPB from *T. b. brucei* and *L. major*, a gene deletion study suggested that OPB is not an essential virulence factor in *L. major* (Munday *et al.*, 2011). This was further confirmed in a study by Moss *et al*. (2015) where it was demonstrated that the activity of *Tbb*OPB was not required for infection.

The OPB peptidase is dimeric, and is comprised of a C-terminal catalytic domain with an α/β hydrolase fold, and a N-terminal regulatory domain with a seven-bladed β-propeller (Fig. 1.7) (Canning *et al.*, 2013). The peptidolytic activity of monomeric OPB is limited to peptide substrates which are smaller than 30 amino acid residues, and has a P_1 site specificity for Arg and Lys (Kanatani *et al.*, 1991).

Peptide hormones, which are rich in basic residues, are favoured substrates of OPB, with cleavage of atrial natriuretic factor (Ndung'u *et al.*, 1992; Troeberg *et al.*, 1996; Morty *et al.*, 2005a), adrenocorticotropic hormone (de Andrade *et al.*, 1998; Tsuji *et al.*, 2004), angiotensins I and II (Nishikata, 1984; Kanatani *et al.*, 1991; Bagarozzi *et al.*,

1998), bradykinin potentiator B (Nishikata, 1984), vasoactive intestinal peptide and substance P (Bagarozzi *et al.*, 1998), neurotensin (Kanatani *et al.*, 1991; Troeberg *et al.*, 1996), reduced (Arg8)/ (Lys8)vasopressin (Troeberg *et al.*, 1996), serum thymic factor (Nishikata, 1984), somatostatin-28, glucagons and dynorphin A (Tsuji *et al.*, 2004) being reported in literature. It has been demonstrated that synthetic substrates, up to 17 residues in length, based on the structure of bradykinin, were hydrolysed by recombinant *Tbb*OPB and *Tcr*OPB (Hemerly *et al.*, 2003).

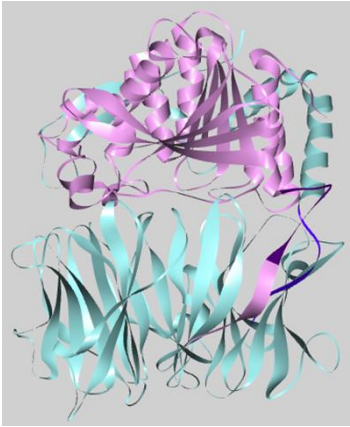


Figure: 1.7: Monomeric 3D structure of *Tbb*OPB. The C-terminal catalytic domain of *Tbb*OPB (PDB: 4BP9) is highlighted in purple, whilst the N-terminal regulatory domain is highlighted in turquoise (Canning *et al.*, 2013).

1.8.2.1 Functions of oligopeptidase B

Invasion of non-phagocytic mammalian cells is mediated by *T. cruzi* metacyclics and trypomastigotes through the action of OPB whereby the free intracellular calcium concentration in host cells is increased (Burleigh and Andrews, 1995). It is thought that OPB acts as a processing enzyme which generates active calcium from a cytosolic precursor molecule (Caler *et al.*, 1998).

It was shown that *Tbb*OPB (Morty *et al.*, 2005b), *Tco*OPB (Pinto Dias, 2006) and *Tev*OPB (Morty *et al.*, 2005a) is released from the respective lysed, dead and dying parasites into the bloodstream (Morty *et al.*, 2005a; Munday *et al.*, 2011), and is, thus, an attractive diagnostic antigen. The OPB from *T. b. brucei* is not inhibited by plasma serpins nor α-2 macroglobulin, and is able to hydrolyse regulatory peptides in the host serum (Troeberg *et al.*, 1996; Morty *et al.*, 2001; Morty *et al.*, 2005a). The action of OPB has been implicated in the disruption of the hormone levels of the infected host (Morty *et al.*, 1999; Morty *et al.*, 2001; Morty *et al.*, 2005a) through the cleavage of peptide hormones, such as atrial natriuretic factor in the blood of *T. b. brucei* infected rats (Tetaert *et al.*, 1993). The disruption of host hormone signalling pathways, through

the action of OPB, is thought to contribute to disease progression (Morty *et al.*, 2005a; Munday *et al.*, 2011).

1.9 Structure-based drug design

The elucidation of the roles that virulence factors play in trypanosomal pathogenesis, together with biochemical characterisation, including structural details, form the basis for the development of specific inhibitors using computational-assisted drug design (Frearson *et al.*, 2007). The availability of large libraries of inhibitor compounds which can be accessed, allows for the testing of novel trypanocides (McKerrow *et al.*, 2008).

Approaches for the discovery of novel chemotherapies include target-based, phenotypic screening and compound repurposing (Field *et al.*, 2017). The most promising methodologies to identify suitable ligands for target proteins are structure-based drug design (SBDD) and virtual ligand screening (VLS) (Bleicher *et al.*, 2003; Foloppe *et al.*, 2006; Klebe, 2006). The use of SBDD provides insight into the interactions between the protein 'receptor' and the ligand, inhibitor or substrate, which highlights chemical modifications required to improve the binding affinity of the ligand to the protein 'receptor'.

A peptidomimetic library based on the substrate P_1 specificity site of *Tbb*MCA2 was synthesised and showed promising results *in vitro* against the recombinant enzyme as well as *T. b. brucei, T. cruzi, L. infantum* and *Plasmodium falciparum* (Berg *et al.*, 2012). When tested *in vitro*, it was suggested that the inhibitors may be acting on multiple targets. Analysis and molecular docking, in an SBDD methodology, is required for the validation of ligand binding to the *Tbb*MCA2 molecular target for the development of possible new ligands targeting the MCAs.

1.10 Antibody production using phage display

Target-specific antibodies, generated using phage display, have applications in proteomics, drug delivery studies and the analysis of intracellular antigens (Hoogenboom, 2005). Phage display is independent from the immune system, whereby *in vitro* selection from combinatorial libraries of immunoglobulin V_H and V_L genes, is used to generate antibodies, which are displayed on phage particles (Smith, 1985; Pini and Bracci, 2000; Schirrmann *et al.*, 2011). The advantage of *in vitro* technologies over animal immunisation is the unlimited and defined source of antigen binders where the DNA is known (Schirrmann *et al.*, 2011).

The process of phage display is based on the linkage of foreign antibody gene fragments to the minor phage coat protein (pIII) of the M13 bacteriophage (Smith, 1985). As a result, the genes encoding the antibody (genotype) is directly linked to the displayed antibody on the phage particle (phenotype). This feature allows the coding sequence of the antibody to be exactly known and exactly reproduced. The antibody::pIII fusion protein are assembled in *E. coli* and secreted into the periplasmic space and are often soluble, stable and folded in their native form (Marston, 1986).

The process of phage display is illustrated in Fig. 1.8 which starts with the construction of a phagemid library whereby the immunoglobulin V_H and V_L genes are amplified and ligated into a phagemid plasmid (Step A). After the propagation of *E. coli* transfected with the phagemid cells (Step B), the scFv::pIII fusion proteins are displayed on the bacterial cell surface (Step C). The phagemids are then rescued by the M13KO7 helper phage which provides the essential proteins (dark grey outline) for the packaging of the recombinant phage DNA within the phage and the display of the fusion proteins on the bacterial cell surface (Webster, 1996) (Step D). As a result, a mosaic population where *E. coli* containing either phagemid or the helper phage exists whereby the WT pIII (dark gray) competes with the scFv::pIII fusion proteins for incorporation into the phage particle (Step E). In Step F, the mosaic population is panned against the immobilised antigen to isolate antigen specific scFv (light blue), whilst non-specific scFvs are discarded (pale yellow). The bound phagemids are eluted from the *E. coli* host, and under antibiotic selection, only the phagemids are transfected back into *E. coli* (Step G). It is at this step, that the WT phagemid is lost. The resulting phagemid population is rescued once again (Step H) and used to pan against the immobilised antigen once again. This process is repeated numerous times to select for scFv with a high affinity for the immobilised antigen.

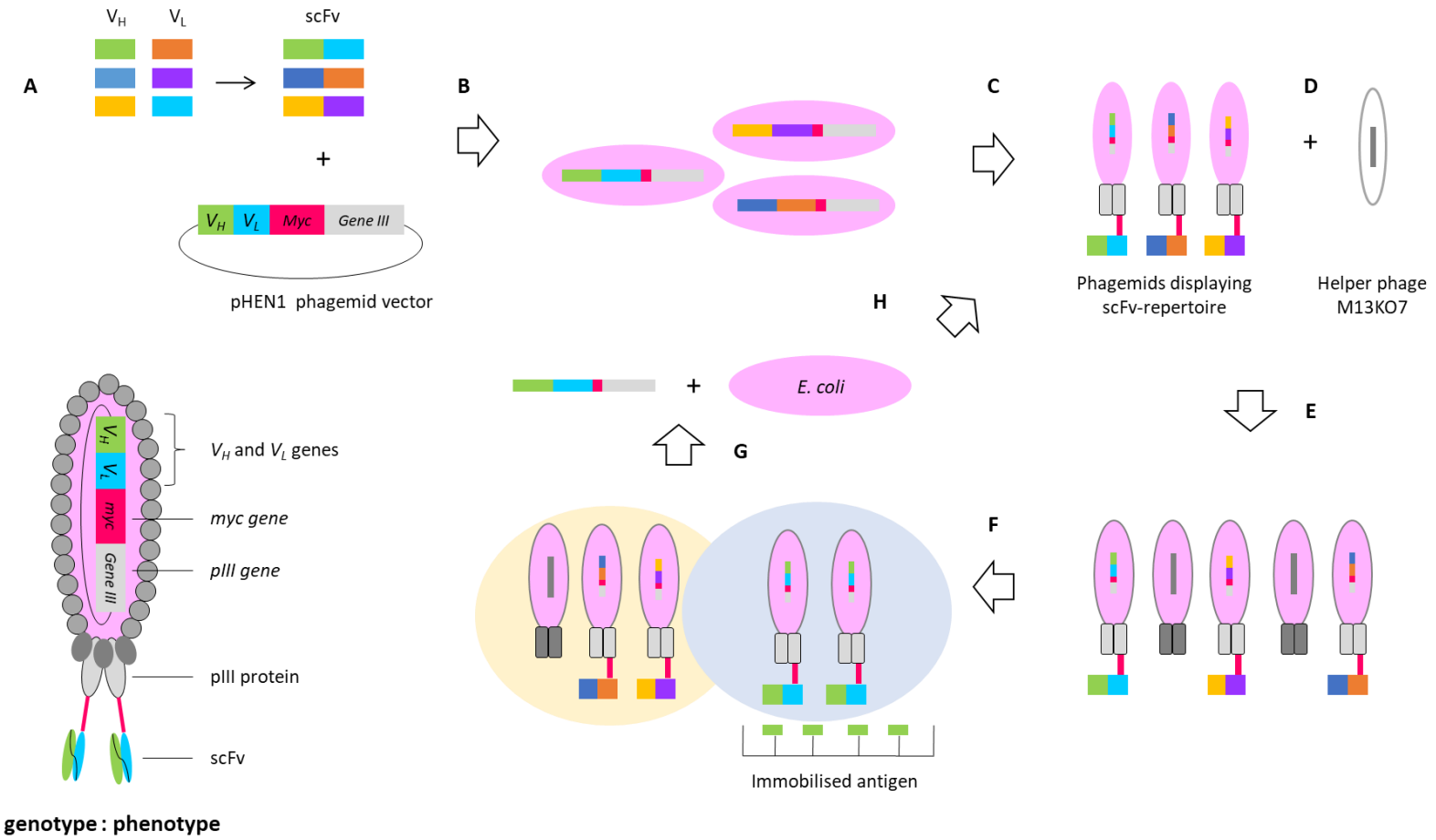


Figure 1. 8: Schematic representation of the genotype:phenotype linkage of scFv and the process of phage display. The resultant myc-tagged scFv antibody fragment::pIII fusion is displayed by the bacterial host (light pink). The genes coding for the immunoglobulin V_H and V_L genes are coloured in light green and turquoise, respectively, and are joined by a linker region $(GGGGS)_3$ which prevents dissociation. A myc sequence, coloured in bright pink, is included to facilitate affinity purification. The pIII phage coat protein is coloured in light grey and the WT pIII in dark grey. The essential packaging proteins, supplied by the helper phage, is shown as the grey outline of the *E. coli*. The process by which the scFv antibodies are selected from the phagemid library is detailed in steps A to H. The stepwise explanation of the process is given in the text. Adapted from McCafferty *et al.* (1990).

1.10.1 Antibody formats produced by phage display

The most common antibody fragments produced by phage display (Fig. 1.9) are the single chain variable fragment (scFv) and the antigen binding fragment (Fab) of all mammalian species as well as the variable heavy domain (V_HH) of camel antibodies (Holt *et al.*, 2003; Hoet *et al.*, 2005; Hust and Dübel, 2005). The monovalent scFv (V_L and V_H), Fab (V_L, C_L, V_H, C_H) and nanobody (V_HH) fragments retain their respective specific antigen binding affinity compared to that of the V_{H1} and C_{H1}, since the antigen binding surface is not altered (Bird *et al.*, 1988; Huston *et al.*, 1988).

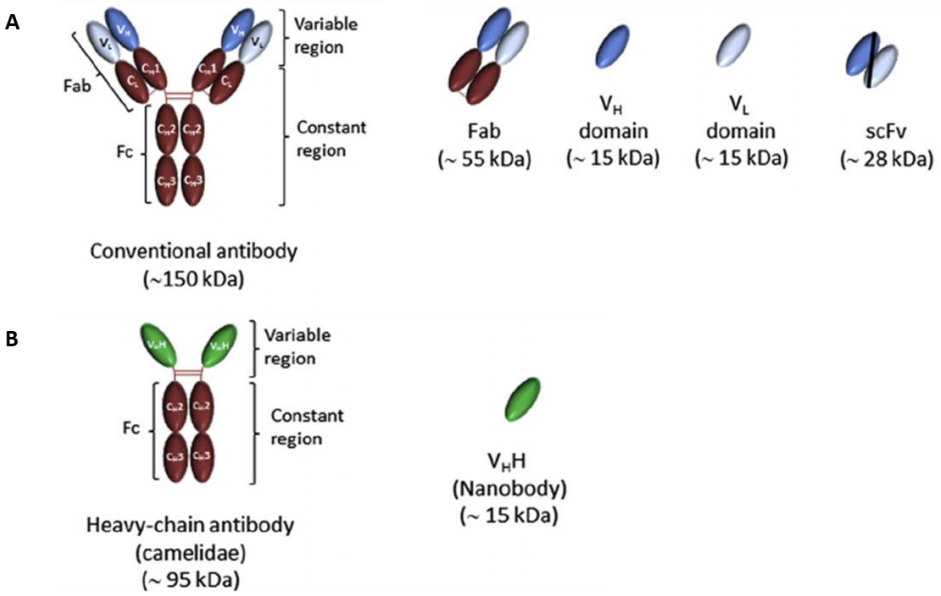


Figure 1.9: Commonly produced antibody fragments using phage display. (Hammarström and Marcotte, 2015).

The evolution of high affinity single V-like domains (V_HH) is an integral part of the camelid immune system (Fig. 1.9, panel B) (De Genst *et al.*, 2005). The V_HH domain possesses extended surface loops which are able to penetrate the narrow cavities on various pathogens' surfaces (Nuttall *et al.*, 2004; Streltsov *et al.*, 2004). In order to escape immunodetection, many pathogens have evolved narrow cavities in their surface antigens in order to make them inaccessible by host antibodies (Janeway Jnr and Medzhitov, 2002). A prime example is evident in the close packing of the VSG dimers on the parasite surface. A V_HH antibody fragment was able to penetrate further between the VSGs on the surface of *T. b. brucei* compared to the Fab antibody fragment (Fig. 1.10) and the anti-VSG IgM antibody (Fig. 1.4) (Stijlemans *et al.*, 2004).

The development of combinatorial libraries for human or mouse V genes is complex due to the requirement for multiple primer sets. In birds, the genes for the heavy (H)

and light (L) chains are subjected to VDJ and VJ rearrangement, respectively. As a result of the incorporation of pseudo V region genes, variability arises by resultant gene conversion (Reynaud *et al.*, 1985; Reynaud *et al.*, 1987; Thompson and Neiman, 1987; Reynaud *et al.*, 1989). This results in the V regions of the chicken immunoglobulins, having identical amino acid residues at both termini. This simplifies the development of a combinatorial library of the naïve chicken antibody repertoire as only one set of primers is required: one for the H and another for the L chain. This characteristic of the genes coding for chicken antibodies was first exploited by Davies *et al*. (1995) using chicken bursal lymphocyte RNA to produce a naïve scFv library, from which scFv antibodies were produced against three proteins. Since chickens are able to produce antibodies against a wide range of antigens (Conroy *et al.*, 2012; Shih *et al.*, 2012), together with the their phylogenetic distance from mammalian species, a combinatorial library using chicken antibody coding genes for the development of diagnostic antibodies would be advantageous (Conroy *et al.*, 2014). In 2004, one such phage display library, using chicken immunoglobulin genes, was developed and named the *Nkuku*® library (van Wyngaardt *et al.*, 2004). The *Nkuku*® phagemid library will be used in the present study.

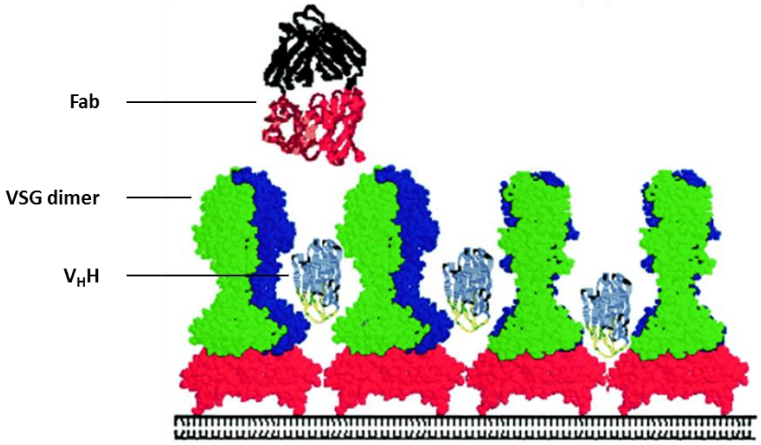


Figure 1.10: Comparison of the penetration of V_HH and Fab antibody fragments in between VSG dimers on trypanosome surfaces. The VSG dimer is coloured green and blue, the V_H, and V_L, C_H and C_L of the Fab fragment are coloured red and black, respectively. The structure of V_HH is coloured in light blue and the complementary determining loops in yellow. (Stijlemans *et al.*, 2004).

1.10.2 Applications of phage display

A study of protein-protein interactions (Hertveldt *et al.*, 2009), identification of immunogenic proteins from pathogens (Stijlemans *et al.*, 2004) and of agonists and antagonists to probe receptor site function (Koolpe *et al.*, 2005) can all be achieved using phage display technology.

In addition, phage display is a cost effective and efficient method to map the epitopes of various antigens which are involved in antibody interaction, thus, providing vital information for the development of diagnostics, immunotherapies and vaccines (Böttger and Böttger, 2009; Wang and Yu, 2009). The *Nkuku*® phage display library was used to map the epitopes of the SAT2 foot-and-mouth disease virus in an attempt to neutralise the virus (Opperman *et al.*, 2012).

Antibodies are used directly in diagnostics for pathogen antigen detection (Dussart *et al.*, 2008) and for competition with serum antibodies for binding to the pathogen antigen (Dong *et al.*, 2007). Single chain variable fragment antibodies can be used in many standard immunodiagnostic tests (Thompson and Neiman, 1987; Nissim *et al.*, 1994; van Wyngaardt and Du Plessis, 1998).

The *Nkuku*® phage display library has been used to generate antibody fragments for the use in ELISA immunodiagnostic tests against bluetongue virus (Fehrsen *et al.*, 2005; Rakabe *et al.*, 2011), African horse sickness virus (van Wyngaardt *et al.*, 2004; van Wyngaardt *et al.*, 2013), a 16 kDa antigen of *Mycobacterium tuberculosis* (Sixholo *et al.*, 2011), and the 65 kDa HSP of *Mycobacterium bovis* (Wemmer *et al.*, 2010). In addition, the HSP65 antibody fragments were engineered to form bivalent constructs, “gallibodies”, which were conjugated to gold particles and used in a sandwich RDT (Wemmer *et al.*, 2010).

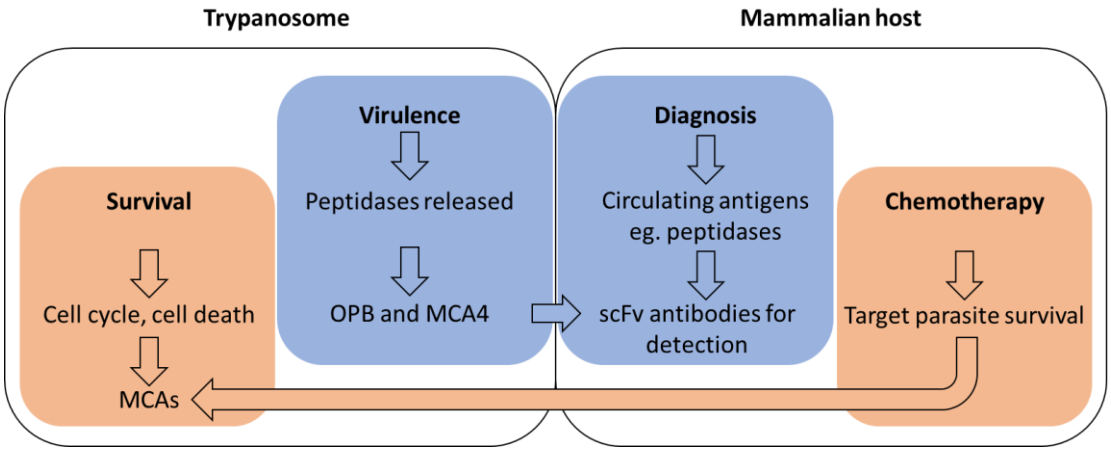
1.11 Objectives of current study

Objective 1: Determination of the applicability of the MCA5s from *T. congolense* and *T. vivax* as a chemotherapeutic target.

1. To recombinantly express and purify MCA2 and -5 in order to raise antibodies for further research.
2. To identify the amino acid residues required for autoprocessing of MCA5.
3. To isolate native MCA5 enzyme.
4. To enzymatically characterise both the recombinant and native MCA5.
5. To perform molecular docking studies of Berg ligands and commercial inhibitors and substrates into the active site of *Tbb*MCA2, for which the library was designed, as well as *Tco*MCA5 and *Tvi*MCA5.

Objective 2: Application of OPB-specific scFv antibody fragments in an antigen detection ELISA format.

1. To pan the *Nkuku*® phagemid library against OPB as a model antigen to select for OPB-specific scFvs.
2. To determine the most likely site at which the selected scFv interacts with OPB.
3. To demonstrate that the isolated scFv detects both recombinant and native OPB in different cellular compartments as well as OPB released into the parasite culture medium.
4. To detect OPB in the sera of experimentally infected cattle using OPB-specific scFv and polyclonal anti-OPB antibodies in an antigen detection sandwich detection ELISA.



CHAPTER 2

CLONING AND EXPRESSION OF THE METACASPASES FROM *TRYPANOSOMA BRUCEI BRUCEI*, MCA2, AND *TRYPANOSOMA VIVAX*, MCA5

2.1 Introduction

African trypanosomosis affects both humans (HAT) and animals (AAT) in rural central Africa. The human infective parasites use wild and domestic animals as a reservoir (Njiokou *et al.*, 2006; Cordon-Obras *et al.*, 2009; Anderson *et al.*, 2011). Thus, the control of AAT is critical for the elimination of HAT, as control of the insect vector is an ineffective strategy (Rotureau and Van Den Abbeele, 2013). The indiscriminate use of trypanocides and the lack of new chemotherapies, resulted in the emergence of drug resistance, and there is consequently a need for new drugs (Field *et al.*, 2017).

Molecular targets for the development of novel chemotherapies are those which are essential for the parasite's survival in the host (Hölzmuller *et al.*, 2008). These targets, also known as virulence factors, are used in an anti-disease strategy rather than an anti-parasite strategy (Antoine-Moussiaux *et al.*, 2009) as the development of a vaccine is unlikely due to antigenic variation (La Greca and Magez, 2011). The anti-disease strategy focusses on targeting of factors which are essential to the growth and survival of the parasite (Stuart *et al.*, 2008).

Homologues of the metazoan caspases are the metacaspases (MCA) which are found in all kingdoms except that of the metazoan (Uren *et al.*, 2000). Apoptosis is a controversial process in unicellular organisms; however, evidence in its favour is mounting (Deponte, 2008; Kaczanowski *et al.*, 2011). Morphological and biochemical features, which are similar to those seen during apoptosis, have been described in *T. brucei* (Welburn *et al.*, 2006b)*, T. cruzi* (Ameisen *et al.*, 1995), *Leishmania* spp. (Gannavaram and Debranbant, 2012)*, Giardia lamblia* and *Plasmodium falciparum* (Bruchhaus *et al.*, 2007). Due to the roles played by caspases in apoptosis and non-apoptotic events, it is thought that the MCAs may function in a manner similar to that of the caspases. As such, the MCAs are considered to be virulence factors as well as attractive drug targets due to their absence in their mammalian hosts.

Despite the conserved caspase-haemoglobinase fold (Aravind and Koonin, 2002), the caspases and MCA are distinctly different in their substrate specificity, activation

mechanisms, calcium dependency and control of peptidolytic activity. In order to determine the role and processes they are involved in, the MCAs need to be identified, functionally characterised and their roles in parasite homeostasis determined.

A number of kinetoplastid MCAs have been studied to date, including the multicopy MCAs, MCA1 (Szallies *et al.*, 2002), MCA2 (Helms *et al.*, 2006; Moss *et al.*, 2007; McLuskey *et al.*, 2012; Machado *et al.*, 2013), MCA3 (Helms *et al.*, 2006), MCA4 (Szallies *et al.*, 2002; Proto *et al.*, 2011) and MCA5 (Helms *et al.*, 2006) from *T. b. brucei*, the MCA3 and -5 from *T. cruzi* (Kosec *et al.*, 2006; Laverrière *et al.*, 2012), and the single copy MCAs from *L. major* (Gonzáles *et al.*, 2007; Zalila *et al.*, 2011; Castanys-Muñoz *et al.*, 2012; Casanova *et al.*, 2015), *L. donovani* (Lee *et al.*, 2007; Raina and Kaur, 2012) and *L. mexicana* (Castanys-Muñoz *et al.*, 2012). Most studies focused on the native MCA function *in vitro* when cell death has been induced in the parasites. Very few studies focus on the characterisation of the recombinant and native enzymes themselves.

Analysis of the phylogenetic relatedness of the multi- and single copy MCA peptidases in *Trypanosoma* spp*. T. cruzi, Leishmania* spp. and the single MCA from *Saccharomyces cerevisiae* (YCA1), shows a clear division between the multi- and single copy groups, shown in yellow and blue, respectively, in Fig. 2.1. Within the multicopy group, a high sequence identity exists between the MCA2 and -3 of *T. b. brucei*, *T. b. gambiense* and *T. evansi*. The MCA2 and -3 of the animal infective *T. congolense* and *T. vivax* species, are not as related to their human infective counterparts. The MCA1, -4 and -5 of *T. b. brucei*, *T. b. gambiense* and *T. evansi* share a 100% sequence identity, and therefore, reference will only be made to those from *T. b. brucei.*

Pseudopeptidases have substitutions for the catalytic residues and have been shown to be inactive (Reynolds and Fischer, 2015). The MCA1s (Fig. 2.1, shown in red) have both the catalytic His and Cys substituted with Tyr and Ser, respectively, whilst the MCA4s (Fig. 2.1, shown in purple) all possess the catalytic His, but have a Ser substitution for the catalytic Cys. The substitution of catalytic residues are common with some pseudopeptidases being shown to play key regulatory roles (Pils and Schultz, 2004; Reynolds and Fischer, 2015). It has been reported that *Tbb*MCA3 processes *Tbb*MCA4, releasing *Tbb*MCA4 which plays a role in cell cycle and parasite virulence (Proto *et al.*, 2011).

The focus of the current study is the MCA5 from the animal infective *T. congolense* and *T. vivax* (Uilenberg and Boyt, 1998). These MCAs differ from the multicopy gene

products, as they possess a long Pro-, Gln-, Tyr-rich C-terminal domain. This extended domain is present in each of the single copy MCAs in both *Trypanosoma* spp. and *Leishmania* spp. (Appendix A3) and is thought to play an important role in protein-protein interactions (Kay *et al.*, 2000). One such example is the apoptotic-like response in *T. cruzi* parasites, in which the *Tcr*MCA5 lacking the C-terminal domain, was overexpressed (Laverrière *et al.*, 2012).

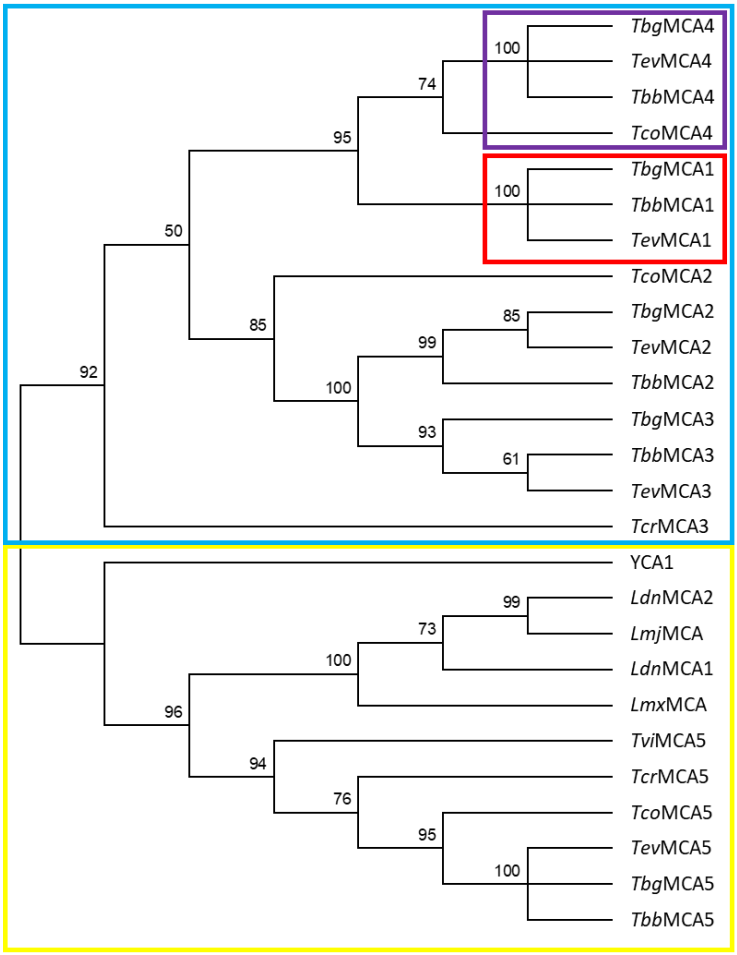


Figure 2.1: Molecular phylogenetic analysis of the kinetoplastid MCAs and the single MCA from *S. cerevisiae*. The bootstrap consensus tree from 500 replicates (Felsenstein, 1985) utilising the maximum likelihood method, was used to deduce the evolutionary history of 26 MCA protein sequences (Jones *et al.*, 1992) with MEGA7 (Kumar *et al.*, 2016). The protein sequences were obtained from TriTrypDB (Aslett *et al.*, 2010); *Tbb*MCA from *T. b. brucei* (927), *Tbg*MCA from *T. b. gambiense* (DAL972), *Tvi*MCA from *T. vivax* (Y486), *Tev*MCA from *T. evansi* (STIB 805), *Tco*MCA from *T. congolense* (IL3000), *Tcr*MCA from *T. cruzi* (Sylvio X10/1), *Lmj*MCA from *L. major* (Friedlin), *Lmx*MCA from *L. mexicana* (MHOM/GT/2001/U1103), *Ldn*MCA from *L. donovani* (BPK282A1) and were compared to the YCA1 from *S. cerevisiae* (UniProt Q08601). The single copy MCAs are grouped in yellow and the multicopy MCAs in blue. The multicopy MCAs with mutations of the catalytic Cys (purple) and both catalytic His and Cys (red) were blocked in their respective colours. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

In the work described in this chapter, the *MCA2* gene from *T. b. brucei* and *MCA5* gene from *T. congolense* were cloned, to include a N-terminal 6xHis tag, and were expressed in the soluble and insoluble fractions using *Escherichia coli*. After solubilisation, on column refolding and purification using nickel affinity chromatography, purified recombinant *Tvi*MCA5 was used to produce antibodies in chickens. The chicken antibodies produced against both *Tco*MCA5 and *Tvi*MCA5 were separately coupled to hydrazide resin to purify the respective MCAs present in the soluble expression fractions.

2.2 Materials and methods

2.2.1 Materials

Molecular biology: The oligonucleotide primers were synthesised at the department of Molecular and Cellular Biology at the University of Cape Town, South Africa. BamHI, EcoRI, NdeI, XhoI [for nomenclature see Roberts *et al.* (2003)], T4 DNA ligase, 10 mM dNTPs, High-fidelity PCR enzyme mix, GeneJet™ Plasmid Miniprep Kit, TransformAid™ Bacterial Transformation Kit, pTZ57R/T cloning plasmid, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (IPTG), isopropyl-β-D-thiogalactopyranoside (X-gal) and dithiothreitol (DTT) were obtained from Fermentas (Vilnius, Lithuania). The pGEM-T® cloning plasmid and T4 DNA ligase were purchased from Promega (Madison, WI, USA) and the pET-28a and pET-32a expression plasmids from Novagen (Darmstadt, Germany). FIREpol® Taq polymerase, 10xPCR reaction buffer and 25 mM $MgCl_2$ were from Solis Biodyne (Tartu, Estonia). The ZymoResearch Clean and Concentrator™ kit and Zymoclean™ Gel DNA Recovery Kit were purchased from Zymo Research (Orange, CA, USA). Seakem®LE agarose was purchased from Lonza (Rockland, ME, USA), ampicillin sodium salt from USB Corporation (Cleveland, OH, USA) and kanamycin from Gibco, (Paisley, UK). Bacteriological agar, tryptone and yeast extract were purchased from Merck Biolab (Darmstadt, Germany). Crystal violet and ethidium bromide were purchased from Sigma (St. Louis, MO, USA). Buffer salts and other common chemicals were purchased from Merck (Germany) and Sigma (St. Louis, MO, USA) and were of the highest purity available. The O'GeneRuler™ 1 kb DNA Ladder from Fermentas (Vilnius, Lithuania) was used in all agarose gels unless otherwise stated.

***E. coli* cells:** *Escherichia coli* cells, JM 109 and BL21 (DE3) strains, were purchased from New England Biolabs (Ipswitch, MA, USA). The JM 109 strain allowed for

blue/white screening for transformants in the presence of IPTG and X-gal. The BL21 (DE3) strain is deficient in both Lon and OmpT peptidase expression (Graslund *et al.*, 2008).

Purification and quantification of recombinant proteins: His-select® nickel affinity resin, 4-chloro-1-naphthol, phosphorylase B (97.4 kDa), bovine serum albumin (BSA, 68 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), soyabean trypsin inhibitor (SBTI, 21.5 kDa), lysozyme (14 kDa) and the semi dry blotter were purchased from Sigma (St. Louis, MO, USA). Equal volumes of 5 mg/ml solutions of phosphorylase B, bovine serum albumin, ovalbumin, carbonic anhydrase, soyabean trypsin inhibitor and lysozyme, made in reducing treatment buffer, were combined to produce the molecular weight marker used in all the SDS-PAGE gels. The BCA™ Protein Assay Kit was purchased from Pierce (Rockford, IL, USA), BioTrace™ nitrocellulose from PALL Corp (Ann Arbor, USA), 4-chloro-1-naphthol and UltraLink® hydrazide resin from ThermoScientific (Waltham, MA USA) and non-fat powdered milk from Amresco (Solon, OH, USA).

Chicken IgY preparation and ELISA: Freund's complete and incomplete adjuvants and bovine serum albumin (BSA, catalogue no.: A7906) were purchased from Sigma (St. Louis, MO. USA). Polyethylene glycol (PEG) M_r 6 000 was purchased from Merck (Darmstadt, Germany), Nunc-Immuno™ Maxisorp 96-well plates from Nunc Intermed (Roskilde, Denmark), 2, 2'-azinobis [3-ethyl-3, dihydrobenzithiazole-6-sulfonate] (ABTS) from Roche (Mannheim, Germany). The BIOTEK® ELx50™ Microplate washer was purchased from BioTek Instruments Inc. (USA) and the FLUORStar Optima Spectrophotometer from BMG Labtech (Offenburg, Germany).

Antibodies: The chicken anti-6xHis antibody was purchased from Merck (Germany) and the rabbit anti-chicken IgY HRPO conjugate from Sigma (St. Louis, MO. USA). Approval was obtained for antibody production using chickens, with protocols approved by the University of KwaZulu-Natal animal research ethics committee (Reference 053/15/Animal).

2.2.2 Cloning of *TbbMCA2* and *TviMCA5* into the pGEM-T® cloning plasmid

The genomic DNA of purified bloodstream trypomastigotes of *T. b. brucei* (strain 927) and *T. vivax* (strain Y486) parasites was isolated utilising the method developed by Medina-Acosta and Cross (1993). Briefly, parasites were washed with PBS [100 mM

Na_2HPO_4, 2 mM KH_2PO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.2, 1 ml], centrifuged (2 000 *g*, 10 min, RT), resuspended in TELT buffer [50 mM Tris-HCl buffer, pH 8.0, 62.5 mM Na_2EDTA, pH 9.0, 2.5 M LiCl, 4% (v/v) Triton X-100, 150 µl] and incubated at RT for 5 min. A phenol-chloroform [1:1 (v/v)] extraction with end-over-end mixing, at RT for 5 min was subsequently performed. To the resulting top phase, 100% (v/v) ethanol (300 µl) was added and mixed using an end-over-end rotator at RT for 5 min to facilitate the precipitation of genomic DNA. Following centrifugation (10 000 *g*, 10 min, RT), the pellet was washed with 100% (v/v) ethanol (1 ml) and any residual ethanol allowed to evaporate at 37 °C for 30 min. The pellet was resuspended in TE buffer [100 mM Tris-HCl buffer, pH 7.5, 10 mM Na_2EDTA, 20 µg/ml RNAse, 30 µl] and incubated at 37 °C for 45 min. A sample of the isolated DNA (2 µl) was electrophoresed on a 1% (w/v) agarose gel, containing ethidium bromide (0.5 µg/ml), in 1xTris-Acetate-EDTA (TAE) buffer (40 mM Tris-HCl buffer, 100 mM acetic acid, 1 mM Na_2EDTA, pH 8.0).

Using Primer 3 (Rozen and Skaletsky, 2000), primers were designed to amplify the 1044 bp and 1 569 bp genes which encodes the MCA2 protein from *T. b. brucei* (strain 927) [GenBank® accession: AAX80349.1] and the MCA5 protein from *T. vivax* (strain Y486) [GenBank® accession: CCC50891.1], respectively. To facilitate subcloning into expression plasmids, restriction sites were included in the designed primers. In Table 2.1, the primers used for PCR amplification of the *TbbMCA2* and the *TviMCA5* genes as well as those required for colony PCR of the recombinant cloning and expression plasmids, are listed. In addition, the cloning and expression plasmids, into which the genes were ligated, are specified (Appendix A4).

The isolated DNA was used as the PCR template. Briefly, the final concentrations of the PCR master mix were: 0.25 µM of each gene primer, 1 x High-fidelity PCR enzyme mix buffer, 2.5 mM $MgCl_2$, 1 U High-fidelity PCR enzyme mix and 0.25 mM dNTPs in a total reaction volume of 50 µl. The PCR amplification of the *TbbMCA2* and *TviMCA5* genes was performed with incubation at 95 °C for 2 min as the initial DNA denaturation step, followed by 30 cycles of 95 °C for 10 s, 55 °C for 15 s and 72 °C for 1 min. A final elongation step was carried out at 72 °C for 7 min. A sample of the amplified gene products (2 µl) was electrophoresed on a 1% (w/v) agarose gel, containing ethidium bromide (0.5 µg/ml), in 1xTAE buffer. The remaining reaction mixture was purified using the Zymo Research Clean and Concentrator™ kit, as per the manufacturer's instructions. A sample of the amplified *TbbMCA2* and *TviMCA5* genes were sequenced

at the Central DNA Facilities, Stellenbosch University, South Africa, using the specific gene primers.

Table 2.1: Primer sequences to be used throughout the *Tbb*MCA2 and *Tvi*MCA5 cloning process.

Primer	5' - 3'	Cloning plasmid	Expression plasmid
TbbMCA2 forward (NdeI)	GCC ATA TGT GCT CCT TAA TTA CAC AAC TCT GTG	pTZ57R/T	pET-28a
TbbMCA2 reverse (XhoI)	CTC GAG CTA TTG GAT AGA TCT GTC AAC AG		
TviMCA5 forward (BamHI)	AAG GAT CCA TGA ATA TTC TTA CC GA TCT CTT TTT G	pGemT®	pET-32a
TviMCA5 reverse (EcoRI)	AAG AAT TCT GTG ATA CAT CAC TTG TGA CCA		
T7 promoter	TAA TAC GAC TCA CTA TAG GG	pGemT®	pET-28a and pET-32a
T7 terminator	CTA GTT ATT GCT CAG CGG TG	pTZ57R/T pGemT®	
SP6 promoter	ATT TAG GTG ACA CTA TAG	pGemT®	-

Underlined sequences correspond to the restriction sites.
Sequences in bold correspond to the start and the stop codon in the forward and reverse primer, respectively.

Ligation of the *TbbMCA2* and *TviMCA5* inserts into the pTZ57R/T and the pGEM®-T cloning plasmids, respectively, was performed using a 3:1 ratio of plasmid to purified PCR product and incubation with 1 U of T4 DNA ligase at 37 °C for 1 h, followed by a further incubation step at RT for 16 h. The ligation mixture was transformed into competent *E. coli* JM 109 cells using the TransformAid™ Bacterial Transformation Kit, as per the manufacturer's instructions. The resulting *E. coli* JM 109 cells (50 µl) from the transformation reaction were plated onto pre-warmed 2xYT plates [1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl, 1% (w/v) bacteriological agar, containing 50 µg/ml ampicillin, 20 µg/ml X-gal, 10 µg/ml IPTG] and incubated at 37 °C for 16 h. The incorporation of X-gal and IPTG onto the 2xYT plates facilitated blue/white screening in order to identify which clones contained the plasmid with the ligated gene insert. The white colonies, which contained the recombinant pTZ57R/T and pGEM®-T plasmid, were selected and grown in 2xYT medium (containing 50 µg/ml ampicillin, 5 ml) at 37 °C for 16 h with agitation. The plasmid DNA of the recombinant colonies was isolated using the GeneJet™ Plasmid Miniprep Kit, as per the manufacturer's instructions. The isolated plasmid DNA was used as the template for colony PCR using the previously described PCR amplification reaction conditions, with the exception that FIREpol® Taq polymerase was used instead of the High-fidelity taq polymerase. The

colony PCR products (5 µl) along with samples of the plasmid DNA (5 µl) were electrophoresed on a 1% (w/v) agarose gel, containing ethidium bromide (0.5 µg/ml), in 1xTAE buffer.

The plasmid DNA of the positive clones of recombinant *TbbMCA2*-pTZ57R/T and *TviMCA5*-pGEMT was sequenced at the Central Analytical Facilities, Stellenbosch University, South Africa using both the gene and plasmid specific primers.

2.2.3 Subcloning of the *TbbMCA2* and *TviMCA5* gene constructs into the bacterial pET-28a and pET-32a expression plasmids

The recombinant *TbbMCA2*-pTZ57R/T and *TviMCA5*-pGEMT plasmids were subjected to a restriction digestion (50 µl) with NdeI and XhoI, BamHI and EcoRI, respectively, in 2xTango buffer at 37 °C for 4 h with subsequent heat deactivation at 80 °C for 15 min. The reaction mixture (50 µl) was electrophoresed on a 1% (w/v) agarose gel in 1xTAE buffer, containing crystal violet (10 µg/ml). Bands were visualised using crystal violet instead of ethidium bromide to avoid DNA damage by ultraviolet light which is required to excise ethidium bromide stained bands (Rand, 1996). The gel was viewed on a light box, the 1 044 bp and 1 569 bp *TbbMCA2* and *TviMCA5* products were excised and subsequently purified using the Zymoclean™ Gel DNA Recovery Kit as per the manufacturer's instructions.

The restriction digestion of the pET-28a plasmid, with NdeI and XhoI, and the pET-32a plasmid, with BamHI and EcoRI, was done in 2xTango buffer at 37 °C for 4 h with subsequent heat deactivation at 80 °C for 15 min. A 3:1 ratio of plasmid to gene insert was incubated with 1 U of T4 DNA ligase at 37 °C for 1 h followed by a further incubation step at RT for 16 h.

The ligation mixture was transformed into *E. coli* BL21 (DE3) cells by $CaCl_2$ transformation (Cohen *et al.*, 1972; Sambrook *et al.*, 2001). Briefly, 2xYT medium (10 ml) was inoculated with a single *E. coli* BL21 (DE3) colony and grown at 37 °C for 16 h with agitation. The overnight culture was diluted 1:100 with fresh 2xYT medium and grown at 37 °C with agitation until an OD_{600} of 0.4 was obtained. The cultures were transferred into ice cold, sterile centrifuge tubes and incubated on ice for 10 min. The cells were pelleted (4 500 *g*, 10 min, 4 °C) and resuspended in ice cold, sterile $CaCl_2$ solution [60 mM $CaCl_2$, 10 mM HEPES, pH 7.0, 40 ml]. The cell solution was pelleted again (4 500 *g*, 10 min, 4 °C) and resuspended in ice cold, sterile $CaCl_2$ solution (2 ml).

Competent *E. coli* cells (20 µl) were incubated with the ligation mixture (1 µl) on ice for 30 min. Thereafter, cells were heat shocked at 42 °C for 90 s and placed on ice immediately for 2 min. The cells were added to pre-warmed super optimal cataboliser with catabolite repression (SOC) medium [2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose, 80 µl] and incubated at 37 °C for 1 h with gentle agitation. The resulting cell mixture (100 µl) was plated onto pre-warmed 2xYT plates containing 34 µg/ml kanamycin for pET-28a and 50 µg/ml ampicillin for pET-32a recombinant clones, and incubated at 37 °C for 16 h.

Recombinant *TbbMCA2*-pET28a and *TviMCA5*-pET32a colonies were grown in 2xYT medium (5 ml), containing 34 µg/ml kanamycin for pET-28a and 50 µg/ml ampicillin for pET-32a, and grown at 37 °C for 16 h with agitation. The plasmid DNA was isolated using the GeneJet™ Plasmid Miniprep Kit, as per the manufacturer's instructions and used as the template for the colony PCR. Using the previously described PCR amplification reaction conditions, with the exception that FIREpol® Taq polymerase was used instead of the High-fidelity Taq polymerase, a colony PCR was performed. The colony PCR products (5 µl) along with samples of the plasmid DNA (5 µl) were electrophoresed on a 1% (w/v) agarose gel, containing ethidium bromide (0.5 µg/ml), in 1xTAE buffer.

2.2.4 Recombinant expression of *Tbb*MCA2 and *Tvi*MCA5

Single colonies of positively identified recombinant *TbbMCA2*-pET28a and *TviMCA5*-pET32a BL21 (DE3) clones were used to inoculate terrific broth (TB) medium [1.2% (w/v) tryptone, 2.4% (w/v) yeast extract, 0.4% (v/v) glycerol, 0.17 M KH_2PO_4, 0.72 M K_2HPO_4, 100 ml] which contained either 34 µg/ml kanamycin for pET-28a or 50 µg/ml ampicillin for pET-32a. After incubation at 37 °C for 16 h in baffled flasks, the cells were pelleted by centrifugation (5 000 *g*, 10 min, 4 °C) and resuspended in 1% (v/v) Triton-X-100-PBS (5 ml). A final concentration of 1 mg/ml lysozyme was added and incubated at 37 °C for 30 min. The cell suspension was frozen at -70 °C for 1 h and subsequently thawed at RT. The cell suspensions were sonicated four times for 30 s each and the cellular debris along with the insoluble inclusion bodies were pelleted from the soluble protein lysate by centrifugation (5 000 *g*, 10 min, 4 °C). The insoluble pellet was resuspended in 1% (v/v) Triton-X-100-PBS (5 ml), was stored at -20 °C.

Samples of the supernatant and the pellet, containing the soluble and insoluble fractions, respectively, were electrophoresed on four 12.5% reducing SDS-PAGE gels (Laemmli, 1970) with one stained with Coomassie Blue R-250 and the others transferred onto nitrocellulose using the semi dry blotter from Sigma with the transfer of proteins confirmed by staining with Ponceau S [0.1% (w/v) Ponceau S, 15% (v/v) acetic acid]. The unoccupied sites on the nitrocellulose membrane were blocked with 5% (w/v) non-fat milk powder in TBS (20 mM Tris-HCl buffer, 200 mM NaCl, pH 7.4) at RT for 1 h. The nitrocellulose was washed three times over 15 min with TBS. Thereafter, chicken anti-6xHis IgY [1:5 000 in 0.5% (w/v) BSA-PBS], chicken anti-*Tco*MCA5 IgY and chicken anti-*Tvi*MCA5 IgY [10 µg/ml in 0.5% (w/v) BSA-PBS] was added and incubated at 4 °C for 16 h. The nitrocellulose was washed and incubated with rabbit anti-chicken IgY HRPO conjugate [1:5 000 in 0.5% (w/v) BSA-PBS] at RT for 1 h. The nitrocellulose was washed before the addition of 4-chloro-1-naphthol·H_2O_2 substrate [0.06% (w/v) 4-chloro-1-naphthol, 0.1% (v/v) methanol and 0.0015% (v/v) H_2O_2 in PBS] and allowed to develop in the dark.

2.2.5 Solubilisation, refolding and purification of recombinant *Tbb*MCA2 and *Tvi*MCA5

Based on the success of the solubilisation and purification of recombinant *Tco*MCA5 (Eyssen, 2013) using the method which was adapted Schlager *et al.* (2012), this method was employed for recombinant *Tbb*MCA2 and *Tvi*MCA5. Sarkosyl solubilisation was performed on the inclusion bodies which were pelleted after expression in TB. The insoluble expression pellet was resuspended in lysis buffer [8 mM Na_2HPO_4, 286 mM NaCl, 1.4 mM KH_2PO_4, 2.6 mM KCl, 1% (w/v) SDS, pH 7.4, 1 mM DTT, 10 ml for a 250 ml culture], sonicated twice over 4 min on ice, and incubated on ice for 30 min. The solubilised proteins were separated from the insoluble debris by centrifugation (13 000 *g*, 20 min, 4 °C).

The refolding and purification of solubilised recombinant *Tbb*MCA2 and *Tvi*MCA5 was performed by immobilised metal affinity chromatography. This purification method will be referred to as nickel affinity purification throughout this manuscript. Briefly, His-select® nickel affinity resin (1 ml) was placed in a 10 ml chromatography column, washed with dH_2O (2 ml) and equilibrated with wash buffer [8 mM Na_2HPO_4, 286 mM NaCl, 1.4 mM KH_2PO_4, 2.6 mM KCl, 0.1 % (w/v) sarkosyl, pH 7.4, 5 ml]. The solubilised protein lysate (5 ml) was incubated with the resin at 4 °C for 3 h with agitation using an end-over-end rotator. The unbound proteins were collected, and the resin was washed

with PCW buffer (50 ml) until an absorbance value at 280 nm of 0.02 was reached. The bound proteins were eluted in 1 ml fractions with elution buffer [8 mM Na_2HPO_4, 286 mM NaCl, 1.4 mM KH_2PO_4, 2.6 mM KCl, 0.1% (w/v) sarkosyl, 50 mM imidazole, pH 7.4, 10 ml]. The column was regenerated using dH_2O (2 ml) followed by 6 M guanidine hydrochloride (5 ml), dH_2O (3 ml) and wash buffer (3 ml) before storage in 30% (v/v) ethanol at 4 °C. Samples of the unbound and the eluted fractions were electrophoresed on a 12.5% reducing SDS-PAGE gel (Laemmli, 1970) and stained with Coomassie Blue R-250. The western blot of *Tvi*MCA5 was performed as per Section 2.2.4 using chicken anti-*Tvi*MCA5 IgY.

The protein concentration was determined using the BCA™ protein assay kit.

2.2.6 Antibody production and ELISA optimisation

2.2.6.1 Preparation of immunogen, the immunisation of chickens and IgY isolation

Ethical clearance for animal procedures was obtained from the University of KwaZulu-Natal animal research ethics committee (Approval number: 053/15/Animal). Chickens were used to raise antibodies against recombinant *Tvi*MCA5. Nickel affinity purified, full length, recombinant *Tvi*MCA5 (50 µg/ml, 1.5 ml) was added to an equal volume of Freund's complete adjuvant and triturated to form a stable water-in-oil emulsion prior to immunisation. Two chickens were immunised intramuscularly on either side of the breast bone. Booster injections, using Freund's incomplete adjuvant, were given at weeks 2, 4 and 6. Eggs collected prior to immunisation served as the pre-immune control.

Chicken immunoglobulin (IgY) was isolated from the yolks of the eggs collected at each week as described by Goldring and Coetzer (2003). Briefly, the egg yolk was separated from the egg white and rinsed with water. The yolk sac was punctured, the yolk collected, and the volume was determined. The yolk was mixed with two yolk volumes of IgY buffer [100 mM NaH_2PO_4, pH 7.6, 0.02% (w/v) NaN_3], 3.5% (w/v) PEG 6 000 was added and dissolved with stirring. The resulting solution was centrifuged (4 420 *g*, 30 min, RT), and the supernatant was filtered through absorbent cotton wool. The final PEG concentration was increased to 12% (w/v) by the addition of 8.5% (w/v) PEG 6 000 and was once again dissolved by stirring. The resulting solution was centrifuged (12 000 *g*, 10 min, RT), and the resulting pellet was dissolved in IgY buffer equal to the yolk volume. Finally, 12% (w/v) PEG 6 000 was added and dissolved with

stirring. The resulting solution was centrifuged (12 000 *g*, 10 min, RT), and the resulting pellet was dissolved in final IgY buffer [100 mM NaH_2PO_4, pH 7.6, 0.1% (w/v) NaN_3], equal to a sixth of the yolk volume, and stored at 4 °C. The concentration of isolated IgY was determined spectrophotometrically at 280 nm using an extinction coefficient of $E^{1\ mg/ml}_{280\ nm}$ = 1.25 (Goldring *et al.*, 2005).

2.2.6.2 ELISA evaluation of antibody production

The progress of anti-*Tvi*MCA5 IgY antibody production during the immunisation period was monitored by using an enzyme-linked immunoadsorbent assay (ELISA). The wells of 96-well Nunc-Immuno™ Maxisorp ELISA plates were coated with recombinant *Tvi*MCA5 (1 µg/ml, 100 µl/well) in PBS for 16 h at 4 °C. The coating solution was discarded, and the plates were blocked with blocking buffer [0.5% (w/v) BSA-PBS, 200 µl/well] to prevent non-specific binding of antibodies, and incubated at 37 °C for 1 h. The wells were washed three times with 0.1% (v/v) Tween-20-PBS (PBS-T) using a BIOTEK® ELx50™ microplate washer and incubated with anti-*Tvi*MCA5 IgY primary antibody diluted in blocking buffer (100 µl/well) at 37 °C for 2 h. The wells were washed and rabbit anti-chicken IgY HRPO conjugate, diluted in blocking buffer (1:5 000, 100 µl/well), was added and incubated at 37 °C for 1 h. The wells were washed and the ABTS substrate solution [0.05% (w/v) ABTS, 0.0015% (v/v) H_2O_2 in 0.15 M citrate-phosphate buffer, pH 5.0, 100 µl/well] was added. The plate was incubated in the dark for 15 min prior to reading the absorbance at 405 nm using the FLUORStar Optima Spectrophotometer for a further three 15 min intervals until absorbance values above 1.0 were reached.

2.2.7 Immunoaffinity purification of soluble, recombinant *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5

2.2.7.1 Antibody coupling to UltraLink® hydrazide resin

Using the UltraLink® hydrazide resin, it is possible to form a hydrazone bond with an oxidised glycoprotein. Using this resin, chicken anti-MCA5 IgYs were immobilised to the UltraLink® hydrazide resin by the oxidation of the carbohydrates on the Fc region and the formation of a hydrazone bond (Fig. 2.2). As a direct consequence of immobilisation at the Fc region, the antibodies are in the correct orientation to bind the specified antigen.

Due to the cleavage of the N-terminal 6xHis tag upon overexpression of soluble recombinant MCAs, nickel affinity purification could not be employed. Thus, immunoaffinity purification was performed by the coupling of chicken anti-*Tco*MCA5 IgY as well as anti-*Tvi*MCA5 IgY antibodies to Ultralink® hydrazide resin. This purification method will be referred to as immunoaffinity purification throughout this manuscript. Briefly, the chicken IgY antibodies (10 mg/ml, 1 ml), which had the highest signal in the ELISA (Appendix A5), was oxidised with 25 mM sodium meta periodate for 30 min at RT and desalted, using the Zebra desalting spin column as per the manufacturer's instructions. The oxidised IgY antibody mixture was agitated with the hydrazide resin (1 ml), which had been previously equilibrated with coupling buffer (100 mM NaH_2PO_4, pH 7.0, 5 ml), using an end-over-end rotator for 2 h at RT followed by a further 16 h at 4 °C. The unbound fraction was collected, and the resin was washed with coupling buffer (10 ml) followed by 1 M NaCl (10 ml) and coupling buffer containing 0.05% (w/v) NaN_3 (10 ml). Coupling efficiency was determined using the BCA™ protein assay kit.

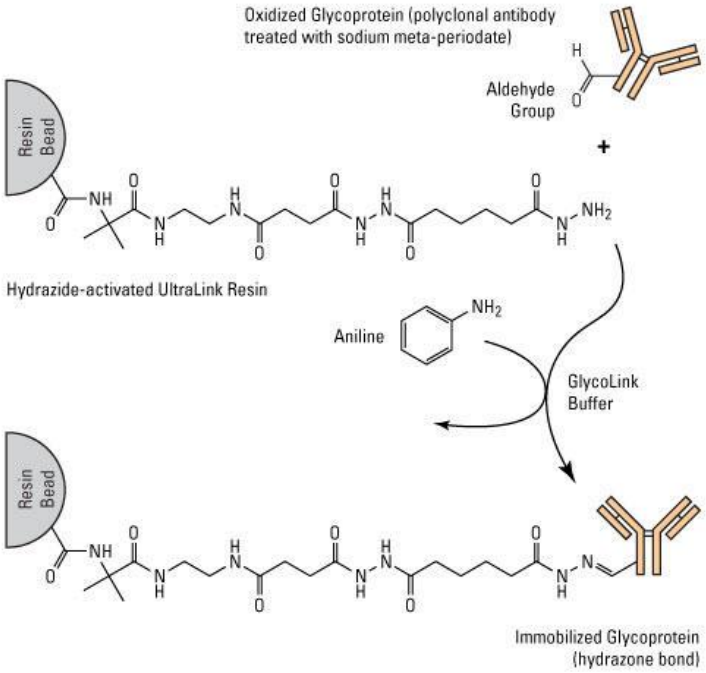


Figure 2.2: Chemistry of the coupling of antibody glycoproteins to UltraLink® hydrazide resin. (ThermoFisher Scientific technical manual #53149).

2.2.7.2 Immunoaffinity purification

The soluble fractions of recombinantly expressed *Tbb*MCA2 and *Tco*MCA5 were purified using the anti-*Tco*MCA5 IgY-hydrazide resin and *Tvi*MCA5 with the anti-*Tvi*MCA5 IgY-hydrazide resin. The resin was incubated with the soluble

expression fraction (8 ml) with agitation using an end-over-end rotator for 3 h at 4 °C. The unbound fraction was collected, and the resin washed with coupling buffer (50 ml) until an absorbance value at 280 nm of 0.02 was obtained. The bound recombinant MCA was eluted with elution buffer (100 mM glycine-HCl, pH 2.8). Fractions (900 µl) were collected into microfuge tubes containing neutralisation buffer (1 M Tris-HCl buffer, pH 8.5, 100 µl) and each was immediately mixed by inversion. The resins were washed with coupling buffer (20 ml) and stored at 4 °C. Samples of the unbound and the eluted fractions were electrophoresed on a 12.5% reducing SDS-PAGE gel (Laemmli, 1970), transferred onto nitrocellulose and probed with chicken anti-*Tco*MCA5 IgY and anti-*Tvi*MCA5 IgY antibodies as described in Section 2.2.4.

2.3 Results

2.3.1 Cloning of the *TbbMCA2* and *TviMCA5* genes into the pTZ57R/T and pGEM®-T cloning plasmids

The single *MCA2* gene from *T. b. brucei* (*TbbMCA2*, GenBank® accession: AAX80349.1), and the single putative full length *MCA5* gene from *T. vivax* (*TviMCA5*, GenBank® accession: CCC5089.1), were identified. Genomic DNA was successfully isolated from *T. b. brucei* (strain 927) and *T. vivax* (strain Y486) bloodstream trypomastigotes (Fig. 2.3, panel A). The isolated genomic DNA of *T. b. brucei* and *T. vivax* were used as the templates for PCR amplification of the 1 044 bp *TbbMCA2* and 1 569 bp*TviMCA5* genes using the specific primers designed for this purpose (Table 2.1). The high-fidelity PCR enzyme mix, comprising a combination of Taq DNA polymerase and a thermostable DNA polymerase with a proofreading ability (ThermoFisher Scientific technical manual K0191), was used for the amplification. These characteristics ensured that the exact sequence would be amplified from the genomic DNA as any transcription errors would result in inaccurate protein translation and, thus, incorrect protein folding in downstream experiments.

The high-fidelity PCR amplification resulted in a single product at the expected sizes of 1 044 bp and 1 569 bp for *TbbMCA2* and *TviMCA5*, respectively (Fig. 2.3, panels B and C). Sequencing of the *TviMCA5* PCR product indicated a perfect match to the DNA sequence found at the GenBank® (TvY486.0907120). However, *TbbMCA2* sequencing results indicated that there were four mismatches when compared to the sequence found at GenBank® (Tb927.6.940). Three of the four mismatches were degenerate. However, the final mutation resulted in the change of Gln71 to an Arg residue. In

relation to the gatekeeper residue (Tyr31) and the catalytic dyad (His158 and Cys213), this mutation is sufficiently distant in the predicted folded protein and should not interfere with any downstream results (Appendix A6).

The PCR products were subsequently purified using the ZymoResearch Clean and Concentrator™ kit. The purified 1 044 bp *TbbMCA2* and 1 569 bp *TviMCA5* PCR products were successfully ligated into the pTZ57R/T and pGEM®-T cloning plasmids, respectively, and upon transformation, non-recombinant blue and recombinant white colonies were obtained after blue/white screening. Plasmid DNA isolation of the recombinant white colonies was performed and used as the template for colony PCR to determine which clones contained the *TbbMCA2* and the *TviMCA5* genes.

Fifteen recombinant *TbbMCA2*-pTZ57R/T colonies were identified by the PCR amplification of the 1 044 bp product using the *TbbMCA2* forward and reverse gene primers (Fig. 2.4, panel A). Using the *TviMCA5* forward and reverse gene primers, the PCR amplification of the 1 569 bp product resulted in the identification of six recombinant *TviMCA5*-pGEMT colonies (Fig. 2.4, panel B).

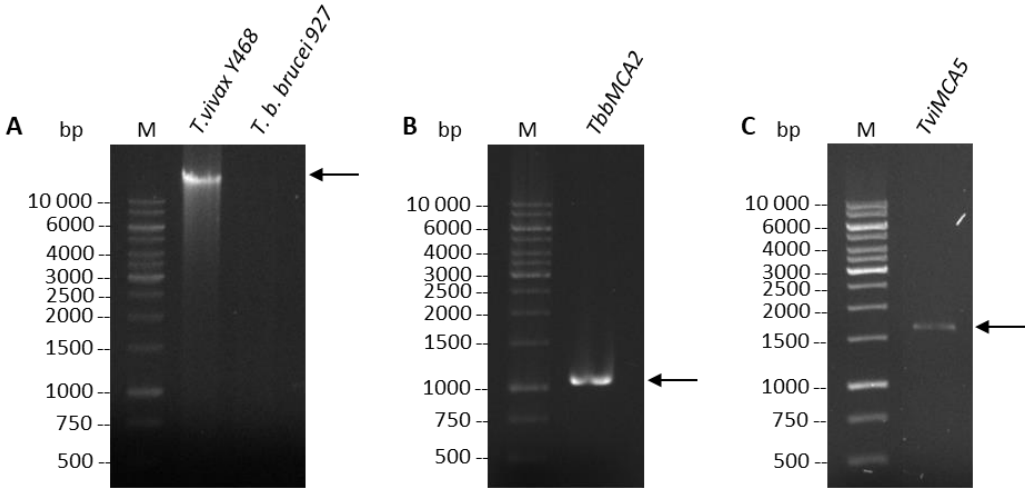


Figure 2.3: Analysis of the isolation of *T. b. brucei* (927) and *T. vivax* (Y486) BSF genomic DNA and PCR amplification of the *TbbMCA2* and *TviMCA5* genes. Samples from **(A)** DNA extraction and high-fidelity PCR amplification of **(B)** *TbbMCA2* and **(C)** *TviMCA5* were electrophoresed on a 1% (w/v) agarose gel containing 0.5 µg/ml ethidium bromide.

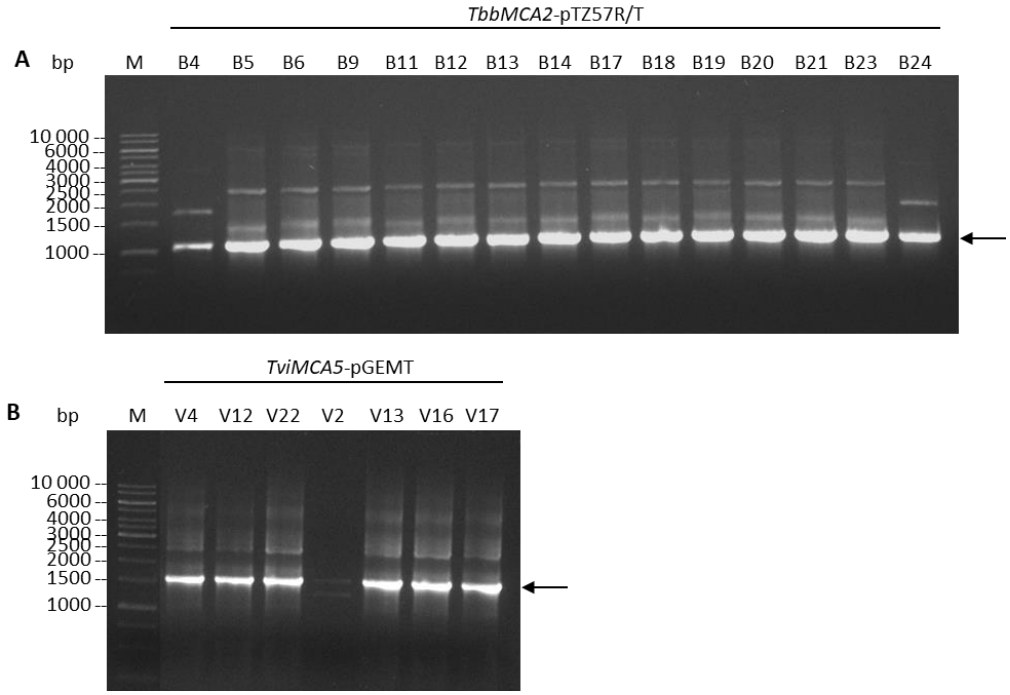


Figure 2.4: Screening for recombinant *TbbMCA2* and *TviMCA5* clones ligated into the pTZ57R/T and pGEM®-T cloning plasmids, respectively, by the PCR amplification of the isolated plasmid DNA. Following the transformation of *TbbMCA2*-pTZ57R/T and *TviMCA5*-pGEMT ligation mixture into competent *E. coli* JM 109 cells, and the subsequent selection of white colonies, the isolated plasmid DNA of **(A)** *TbbMCA2*-pTZ57R/T and **(B)** *TviMCA5*-pGEMT was subjected to PCR using their respective gene forward and reverse primers. Samples were electrophoresed on a 1% (w/v) agarose gel containing 0.5 µg/ml ethidium bromide.

2.3.2 Subcloning of the *TbbMCA2* and *TviMCA5* genes into the bacterial pET-28a and pET-32a expression plasmids

A small scale restriction digestion was performed on the plasmid DNA of the PCR positive recombinant clones. Each of the digestion reactions resulted in the appropriate sized gene insert and plasmid for both *TbbMCA2*-pTZ57R/T and *TviMCA5*-pGEMT. Sequencing of a selection of clones resulted in the same result as obtained when the PCR amplified product was sequenced. The clones selected for sub-cloning were and will now be referred to as B6 from *TbbMCA2*-pTZ57R/T, V4 and V16 from *TviMCA5*-pGEMT, which corresponds to the extracted plasmid DNA.

The isolated plasmid DNA from clones, B6, V4 and V16, were subjected to a large scale restriction digestion to obtain the insert for the subsequent ligation into the expression plasmids. As shown in Fig. 2.5, panel A, after the restriction digestion of B6 and the pET-28a expression plasmid with NdeI and XhoI, two bands were produced at 1 044 bp and 2 886 bp which corresponds to the *TbbMCA2* insert and the pTZ57R/T

cloning plasmid, respectively. The pET-28a expression plasmid was effectively linearised at 5 369 bp. The restriction digestion of V4 and V16 along with the pET-32a expression plasmid, using BamHI and EcoRI, resulted in two bands at 1 569 bp and 3 000 bp corresponding to the *TviMCA5* insert and pGEM®-T cloning plasmid (Fig. 2.5, panel C). The band at 5 900 bp indicated that the pET-32a expression plasmid had been linearised. The restriction mixtures of the B6, V4 and V16 clones were electrophoresed on a 1% (w/v) agarose gel containing 10 µg/ml crystal violet, and the 1 044 bp *TbbMCA2* and the 1 569 bp *TviMCA5* inserts were excised and purified (Fig. 2.5, panel B and D, respectively). The purity, size and approximate concentrations of the isolated *TbbMCA2* and *TviMCA5* inserts along with the linearised pET-28a and pET-32a expression plasmids were determined from Fig. 2.5 panels B and D using the SYNGENE GeneSys image acquisition software.

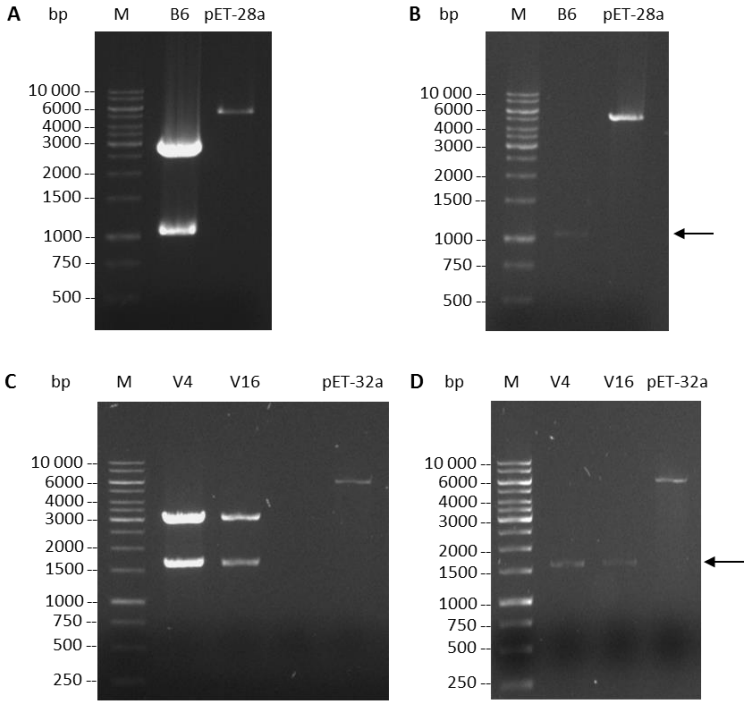


Figure 2.5: Products of gel extraction of the *TbbMCA2* and *TviMCA5* inserts after restriction digestion of positive recombinant *TbbMCA2*-pTZ57R/T clone 6 and *TviMCA5*-pGEMT clones 4 and 16. Restriction digestion was performed on the **(A)** *TbbMCA2*-pTZ57R/T and pET28a plasmid DNA with NdeI and XhoI, **(C)** *TviMCA5*-pGEMT and pET32a plasmid DNA with BamHI and EcoRI for 4 h at 37 °C. The **(B)** *TbMCA2* and **(D)** *TvMCA5* inserts were visualised using 10 µg/ml crystal violet, excised from the 1% (w/v) agarose gel and purified using the E.Z.N.A® gel extraction kit. Samples were electrophoresed on a 1% (w/v) agarose gel containing 0.5 µg/ml ethidium bromide.

Ligation mixtures of the *TbbMCA2* and *TviMCA5* inserts along with the pET-28a and pET-32a expression plasmids, respectively, were prepared and transformed into *E. coli* BL21 (DE3) cells by $CaCl_2$ transformation. Recombinant colonies were obtained for the

B6pET28a ligation, resulting in 4 colony forming units, as well as for the V4pET32a and V16pET32a ligations, with each yielding 2 and 6 colony forming units, respectively.

The plasmid DNA of the resulting recombinant B6pET28a, V4pET32a and V16pET32a clones was isolated and used as the template for colony PCR to amplify the *TbbMCA2* and *TviMCA5* genes using the respective forward and reverse gene primers (Fig. 2.6). In panel A, a band at approximately 1 044 bp was observed for each of the B6-pET28a clones and in panel B, a band at approximately 1 569 bp was observed for V4pET28a clones 1 and 2 and for V16pET32a clones 2, 3, 5 and 6.

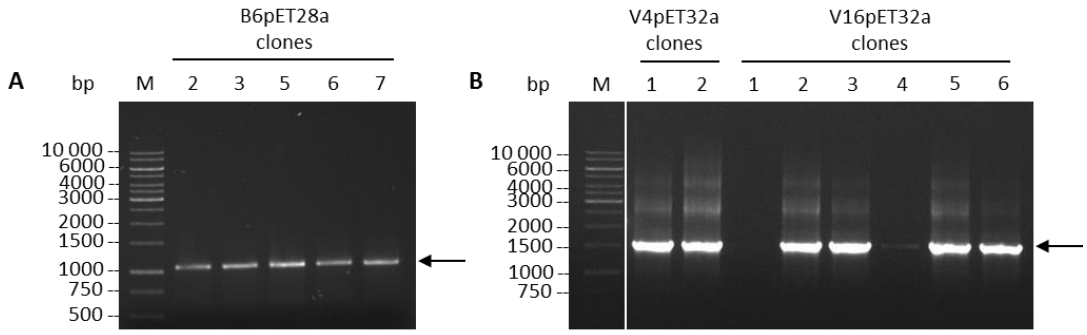


Figure 2.6: Screening for recombinant *TbbMCA2* and *TviMCA5* clones ligated into the pET-28a and pET-32a expression plasmids, from the isolated plasmid DNA by PCR amplification. Colonies were selected following transformation of the **(A)** *TbbMCA2*-pET28a and **(B)** *TviMCA5*-pET32a ligation mixtures into *E. coli* BL21 (DE3) cells. The isolated plasmid DNA from each clone was subjected to PCR amplification using the respective forward and reverse gene primers. Samples were electrophoresed on a 1% (w/v) agarose gel containing 0.5 µg/ml ethidium bromide. M: O'GeneRuler™.

2.3.3 Recombinant expression, solubilisation, refolding and purification of *Tbb*MCA2 and *Tvi*MCA5

The single full length *TbbMCA2* gene (UniProtKB ID: Q585F4) codes for a protein of 347 amino acids in length with an expected molecular size of 37.8 kDa and a p*I* of 5.81 as predicted by the Compute p*I*/Mw tool on the ExPASy server (Gasteiger *et al.*, 2005). The single putative full length *TviMCA5* gene (UniProtKB ID: G0U3N4) codes for a protein of 522 amino acids in length with an expected molecular size of 58.4 kDa and a p*I* of 8.76 as predicted by the Compute p*I*/Mw tool on the ExPASy server (Gasteiger *et al.*, 2005).

2.3.3.1 Expression of recombinant *Tbb*MCA2 and *Tvi*MCA5

The primary structure of a protein plays a pivotal role in its overall solubility (Smialowski *et al.*, 2006). When overexpressed in *E. coli*, proteins containing stretches of 20 or

more hydrophobic residues and comprising of more than 16% aromatic residues (Bertone *et al.*, 2001) were found to be more likely to be insoluble (Christendat *et al.*, 2000). The protein sequences of *Tbb*MCA2 and *Tvi*MCA5 were analysed using the PROSO II sequence-based solubility prediction calculator. It was found that *Tbb*MCA2 had a 38.3% chance of being insoluble when overexpressed in *E. coli*, and *Tvi*MCA5, a 48% chance (Smialowski *et al.*, 2006). The MCA-3 and -5 from *T. cruzi* (Kosec *et al.*, 2006) as well as *Tco*MCA5 (Eyssen, 2013) have PROSO II predictors between 47 and 61% and were reported to be insoluble when overexpressed in *E. coli*. The MCA5s of *T. congolense, T. cruzi* and *T. vivax* share a high sequence identity, between 56 and 67%. Given this, it is probable the *Tvi*MCA5 will be insoluble when overexpressed. The MCA2 differs from the MCAs by its lack of the Pro-, Gln- and Tyr-rich C-terminal extension. But given the fact that the single copy *Tcr*MCA3 was insoluble, there is a chance that *Tbb*MCA2 may also be insoluble when overexpressed.

The MCAs have three domains; namely, the N-terminal domain, the catalytic domain and C-terminal domain. An additional domain was created by the incorporated sequence from the bacterial expression plasmids, comprising of the N-terminal 6xHis tag and sequence of the multiple cloning site up to the restriction site used. These are illustrated for recombinant *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5 in Fig. 2.7 where the estimated sizes of the domains are indicated.

Protein	Vector+ 6xHis	N-terminal domain	Catalytic domain	C-terminal domain
*Tbb*MCA2 42.6 kDa	4.8 kDa	6.8 kDa	27.3 kDa	3.7 kDa
*Tco*MCA5 65.3 kDa	6.2 kDa	6.8 kDa	28.8 kDa	23.5 kDa
*Tvi*MCA5 77.1 kDa	18.6 kDa	7.1 kDa	28.5 kDa	22. 9 kDa

Figure 2.7: Schematic representation of the size and domain composition of the recombinantly expressed *Tbb*MCA2, *Tco*MCA5 in pET-28a and *Tvi*MCA5 in pET-32a. The sizes and regions were annotated using NCBI Conserved Domains Database (Marchler-Bauer *et al.*, 2017).

A 50 to 51% sequence identity is shared between *Tbb*MCA2 and the MCA5s of *T. congolense* and *T. vivax*. As such, chicken anti-*Tco*MCA5 IgY was expected to detect the recombinantly expressed *Tbb*MCA2. Autoprocessing of MCAs generally functions to remove the N- and C-terminal domains from the catalytic domain and is suggested to be required for optimal activity (Gonzáles *et al.*, 2007). Autoprocessing was evident when both recombinant *Tbb*MCA2 and *Tco*MCA5 were overexpressed in *E. coli* (Fig 2.8 and 2.9). Since the sites of autoprocessing are not conserved within the kinetoplastid MCAs (Appendix A3), the allocation of domains to autoprocessed protein fragments of *Tvi*MCA5 were predicted using literature (Vercammen *et al.*, 2004; Gonzáles *et al.*, 2007; Lee *et al.*, 2007; Watanabe and Lam, 2011) as is shown in Fig. 2.9. The sizes of the anti-MCA IgY detected proteins correlate strongly to the estimated molecular weights in Fig 2.7.

Using the chicken anti-*Tco*MCA5 IgY, multiple proteins in both the soluble and insoluble fractions of *Tbb*MCA2 were detected (Fig. 2.8, panel B). The proteins at 82, 72.3, and 68.5 kDa are of a higher molecular weight than the predicted 43 kDa 6xHis N-terminal tagged *Tbb*MCA2. Due to the large size discrepancy between the predicted and actual, no domain allocations could be made with confidence. The 36 and 30 kDa proteins are similar to that of recombinant *Tco*MCA5, which were referred to as lower molecular weight products after autoprocessing (Eyssen, 2013). Given the sizes of the proteins and the estimations made in Fig. 2.7, the 30 kDa may be the catalytic domain and the 36 kDa the catalytic domain before the removal of the N-terminal domain. The 21.5 kDa protein band may be as a result of the partial cleavage of the catalytic domain, once released from the full length protein. Despite numerous attempts, a blot using the anti-6xHis antibody was unsuccessful. Moss *et al.* (2007) showed that mutation of Lys55, interrupted the cleavage of the N-terminal 6xHis tag which facilitated purification of recombinant *Tbb*MCA2 using the soluble expression fraction. However, the presence of *Tbb*MCA2 in the inclusion bodies was not demonstrated in their study. Seen as each of the clones produced similar protein bands in the soluble and insoluble expression fraction, the B6pET28 clone 7 was selected for large scale expression.

The predicted size of *Tvi*MCA5 is 58.4 kDa. However, when factoring in the N-terminal Trx and 6xHis tags incorporated by the expressed vector, recombinant *Tvi*MCA5 has a predicted size of 77.1 kDa (Fig. 2.7). The anti-6xHis antibody detected a single protein at 77 kDa in the insoluble fraction (Fig. 2.9, panel B). This protein was also detected using the chicken anti-*Tco*MCA5 IgY along with other lower molecular weight proteins at 66, 38, 29, and 21.5 kDa in the soluble and insoluble fractions (Fig. 2.9,

panel C). This result is in line with what was seen for recombinant *Tco*MCA5 (Eyssen, 2013). The 77 kDa protein may be as the result of the cleavage of the N-terminal 6xHis tag, followed by the complete removal of the additional sequence from the expression plasmid, to produce the 66 kDa protein. Removal of the C-terminal domain (21.5 kDa band) from the 77 kDa full length protein would produce the 38 kDa protein. Subsequent cleavage of the N-terminal domain of the 38 kDa protein would result in the 29 kDa protein. Seen as each of the clones produced similar protein bands in the soluble and insoluble expression fraction, the V4pET32 clone 2 was selected for large scale expression.

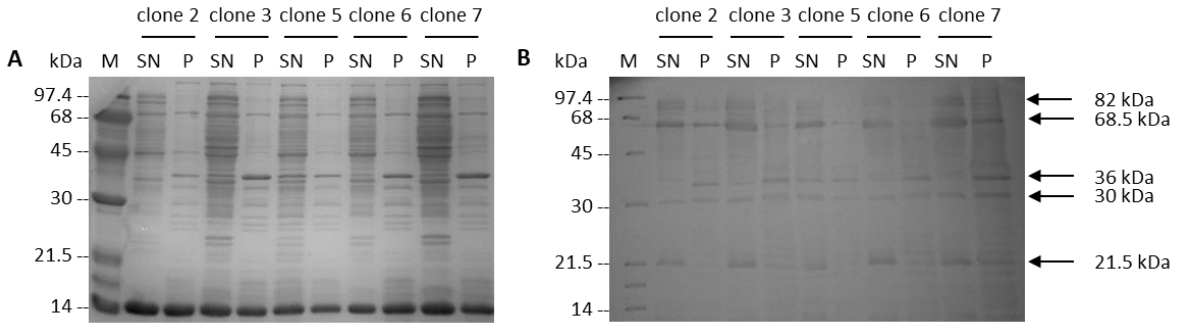


Figure 2.8: Analysis of recombinantly expressed *Tbb*MCA2 in pET-28a. Samples of the soluble (SN) and insoluble (P) fractions of the expression of recombinant *Tbb*MCA2, from the B6pET28a clones 2, 3, 5 and 7 were electrophoresed on two 12.5% reducing SDS-PAGE gels with one **(A)** stained with Coomassie Blue R-250 and the other transferred onto nitrocellulose, blocked with 5% (w/v) milk-TBS and incubated with **(B)** chicken anti-*Tco*MCA5 IgY [10 µg/ml in 0.5% (w/v) BSA-PBS]. Rabbit anti-chicken IgY HRPO conjugate and [1:5 000 in 0.5% (w/v) BSA-PBS] and 4-chloro-1-naphthol·H_2O_2 were used as the detection system.

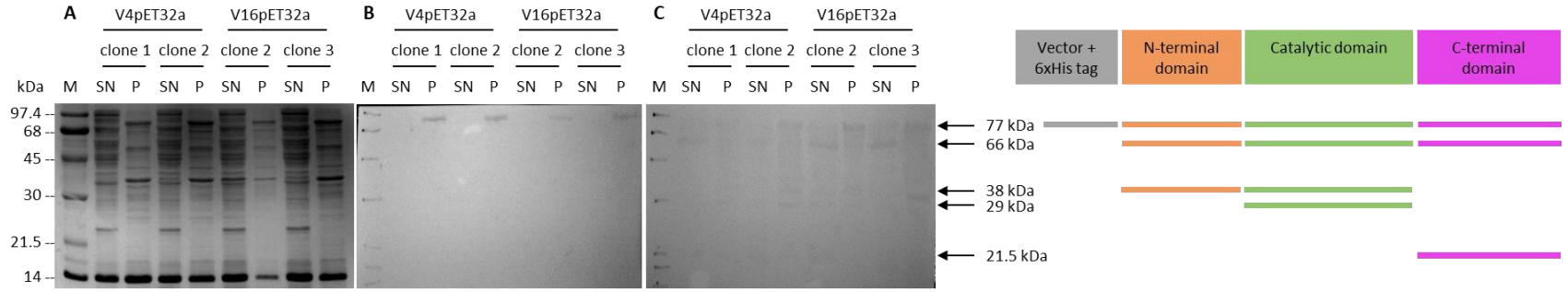


Figure 2.9: Analysis of recombinantly expressed *Tvi*MCA5 in pET-32a. Samples of the soluble (SN) and insoluble (P) fractions of the expression of recombinant *Tvi*MCA5, from the V4pET32a clones 1 and 2, V16pET32a clones 2 and 3, were electrophoresed on three 12.5% reducing SDS-PAGE gels with one **(A)** stained with Coomassie Blue R-250 and the other two transferred onto nitrocellulose, blocked with 5% (w/v) milk-TBS and incubated with **(B)** chicken anti-His IgY [1:5 000 in 0.5% (w/v) BSA-PBS] and **(C)** chicken anti-*Tvi*MCA5 IgY [10 µg/ml in 0.5% (w/v) BSA-PBS]. Rabbit anti-chicken IgY HRPO conjugate and [1:5 000 in 0.5% (w/v) BSA-PBS] and 4-chloro-1-naphthol·H_2O_2 were used as the detection system.

2.3.4 Solubilisation, refolding and purification of recombinantly expressed *Tbb*MCA2 and *Tvi*MCA5

The structure of proteins which have been expressed within inclusion bodies can vary from being completely misfolded to mostly native protein (Bowden *et al.*, 1991; Ventura and Villaverde, 2006). The presence of the N-terminal 6xHis tag in the insoluble fraction and absence in the soluble fraction, suggest that the expressed recombinant *Tbb*MCA2 and *Tvi*MCA5 may be in a partially native state and has retained some degree of activity. This cleavage which is observed in Fig. 2.8 and 2.9 can only be as a result of autoprocessing as peptidase deficient DE3 *E. coli* cells were used, and the addition of bacterial inhibitor cocktail resulted in the same cleavage pattern (results not shown). Solubilisation of the inclusion bodies from the insoluble fraction followed by on column refolding and purification were performed using the sarkosyl method together with nickel affinity purification.

Due to the failure of the anti-6xHis antibody to detect *Tbb*MCA2 in either the soluble or insoluble fraction, both fractions were subjected to nickel affinity purification. No proteins from the soluble fraction were bound to the resin (Fig. 2.10, panel A) when compared to those in panel B which had been solubilised and refolded. Prominent protein bands at 70.2, 32.6 and 29.2 kDa and a faint band at 51.9 kDa were eluted. This confirms that the recombinant *Tbb*MCA2 autopeptidolytically cleaves the N-terminal 6xHis tag in the soluble expression fraction.

When solubilised recombinant *Tvi*MCA5 was purified using nickel affinity resin (Fig. 2.11), numerous proteins, at various sizes, were eluted, which differs from the five proteins from recombinant expression detected by chicken anti-*Tvi*MCA5 IgY (Fig. 2.9, panel C). The approximate molecular weights of the prominent protein bands were 70, 68, 38, 31, 25 and 23 kDa (Fig. 2.11, panel A) of which only 25 and 23 kDa proteins were not detected using the chicken anti-*Tvi*MCA5 IgY (panel B). Previous studies have shown that some MCAs require calcium for autoprocessing (Bozhkov, 2005; He *et al.*, 2008; Ojha *et al.*, 2010; Meslin *et al.*, 2011; Proto *et al.*, 2011; Laverrière *et al.*, 2012; Wen *et al.*, 2013). A possible explanation for the multiple protein bands might that be the nickel ions are functioning in the same capacity as calcium, which assists in autoprocessing. This hypothesis will be referred to as nickel-induced over autoprocessing and will be discussed further in Chapter 3.

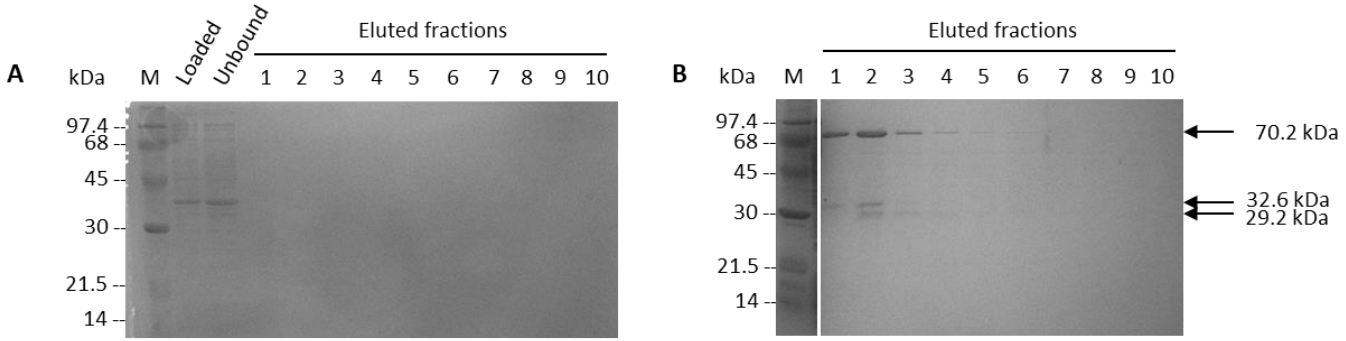


Figure 2.10: Nickel purified recombinant *Tbb*MCA2 from expression supernatant and sarkosyl solubilised inclusion bodies. Samples of the *Tbb*MCA2 nickel purified fractions from the **(A)** soluble expression fraction and **(B)** solubilised and refolded inclusion bodies using the sarkosyl method were electrophoresed a 12.5% reducing SDS-PAGE gel and stained with Coomassie Blue R-250.

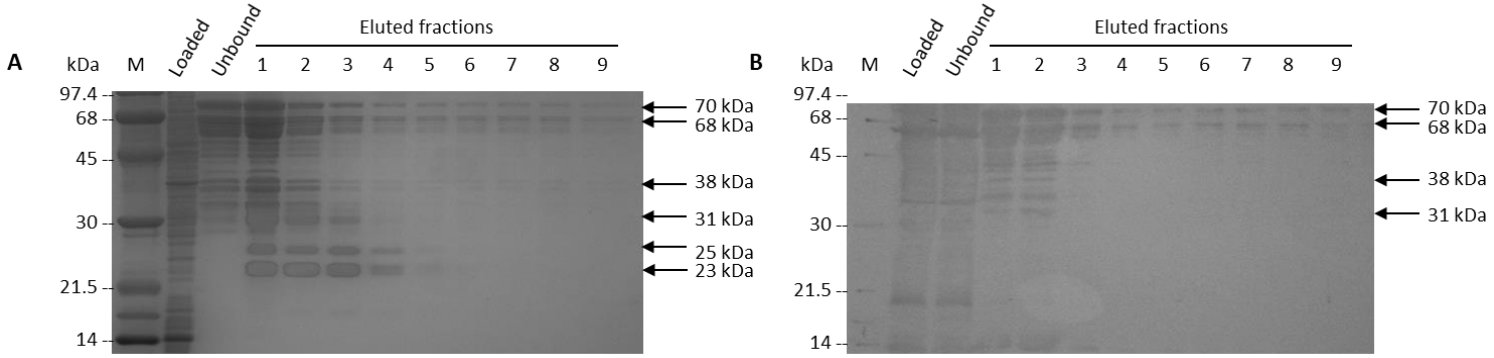


Figure 2.11: Solubilised, refolded and nickel purified recombinant *Tvi*MCA5. Samples of the eluted fractions from the purification of solubilised and refolded recombinant *Tvi*MCA5 from inclusion bodies were electrophoresed on two 12.5% reducing SDS-PAGE gels with one **(A)** stained with Coomassie Blue R-250 and the other transferred onto nitrocellulose, blocked with 5% (w/v) milk-TBS and incubated with **(B)** chicken anti-*Tvi*MCA5 IgY [10 µg/ml in 0.5% (w/v) BSA-PBS]. Rabbit anti-chicken IgY HRPO conjugate [1:5 000 in 0.5% (w/v) BSA-PBS] and 4-chloro-1-naphthol·H_2O_2 were used as the detection system.

2.3.5 Evaluation of anti-*Tvi*MCA5 IgY antibody production by ELISA

The chicken anti-*Tvi*MCA5 IgY antibodies isolated during the immunisation period were used for an initial ELISA to determine when the production of chicken anti-*Tvi*MCA5 IgY antibodies peaked (Appendix A5). The antibodies from chicken 1 at week 9 were used to couple to the UltraLink® hydrazide resin as well as for the determination of the optimal concentration for use in downstream applications.

2.3.6 Immunoaffinity purification of soluble *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5

Since the recombinant MCAs are able to autoprocess thereby cleaving off their N-terminal 6xHis tag in the soluble fraction, an alternative purification strategy was employed. Chicken anti-MCA5 IgY coupled for the hydrazide resin were used in an immunoaffinity purification strategy. The soluble expression fraction of recombinant *Tbb*MCA2 and*Tco*MCA5 was purified using the anti-*Tco*MCA5 IgY immunoaffinity resin, and recombinant *Tvi*MCA5 using anti-*Tvi*MCA5 IgY immunoaffinity resin.

Purification of soluble recombinant *Tbb*MCA2 using the immunoaffinity resin, resulted in the elution of 75.3 and 69.4 kDa proteins in fraction 1, 69.4 kDa in fraction 2, and 56.6 kDa proteins in fractions 6 to10 (Fig. 2.12, panel A). Since there is a large molecular weight discrepancy between the predicted and actual molecular weight of *Tbb*MCA2, as seen in Fig. 2.7 and 2.8, the assignment of most likely domains corresponding to each domain cannot be made with any degree of certainty.

In panel B of Fig. 2.12 it is shown that the purification of soluble *Tco*MCA5 resulted in a 66.8 kDa protein eluted in fractions 1 to 7 together with a 59.1 kDa protein band in fractions 1 and 2. The 66.8 kDa protein may be the full length protein after the cleavage of the N-terminal 6xHis tag. The 59.1 kDa protein may be as a result of the cleavage of the entire N-terminal domain from the full length protein. Cleavage of the C-terminal domain from the full length protein may be the result of the 37.5 kDa protein which was eluted in fractions 3 and 4.

Immunoaffinity purification of soluble*Tvi*MCA5 resulted in the elution of a 78.3 kDa protein band in fractions 1 to 6 which may be as a result of the cleavage of the N-terminal 6xHis tag (Fig. 2.12, panel C). In fractions 3 and 4, two proteins were eluted at 66.4 and 43.9 kDa. The 66.4 kDa protein may be as a result of the cleavage of the

additional sequence from the expression plasmid (Fig. 2.7). The cleavage of the N-terminal domain from the 66.4 kDa protein may result in the 43.9 kDa protein comprised of the catalytic domain and the C-terminal domain.

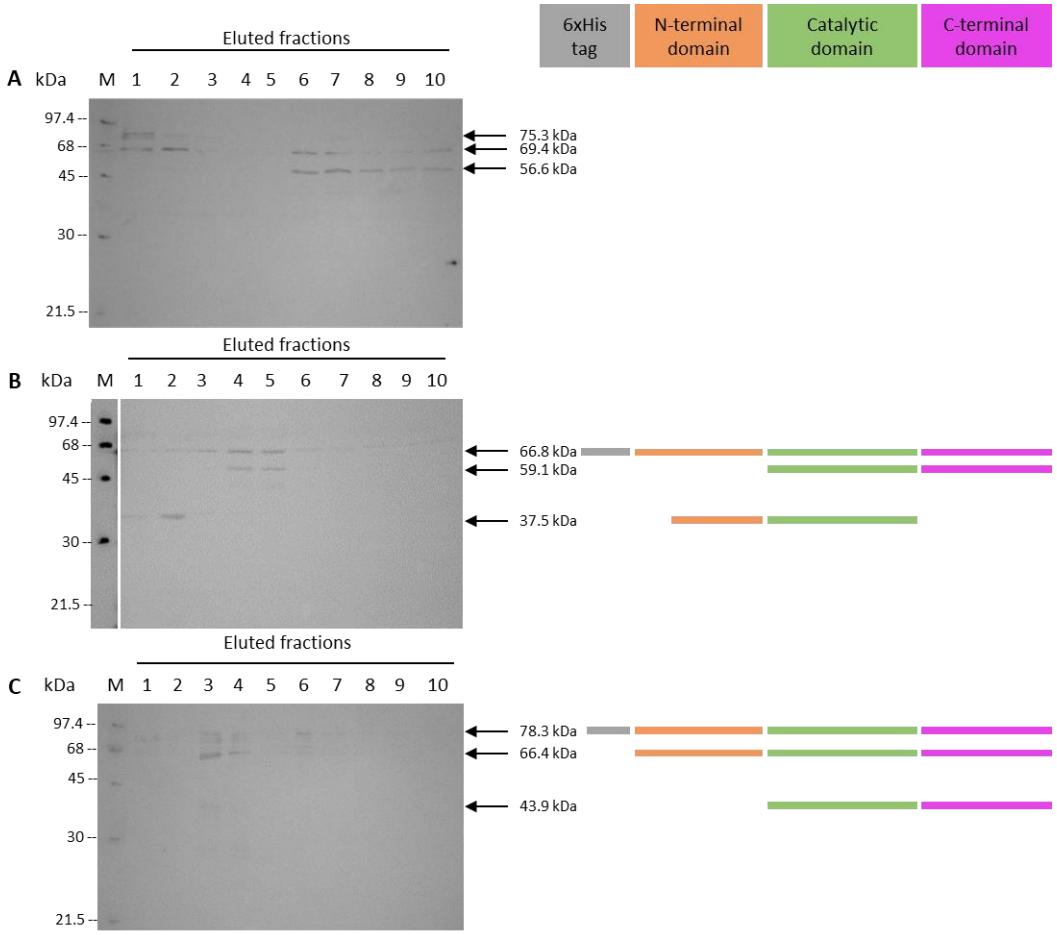


Figure 2.12: Immunoaffinity purification of soluble recombinant *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5. Samples of the eluted fractions from the immunoaffinity purification using chicken anti-MCA5 IgY coupled to hydrazide resins of the soluble recombinantly expressed **(A)** *Tbb*MCA2, **(B)** *Tco*MCA5 and **(C)** *Tvi*MCA5 were electrophoresed on 12.5% reducing SDS-PAGE gels. After transfer onto nitrocellulose and blocked with 5% (w/v) non-fat milk-TBS, the blots were incubated with **(A and B)** chicken anti-*Tco*MCA5 IgY [10 µg/ml in 0.5% (w/v) BSA-PBS] and **(C)** chicken anti-*Tvi*MCA5 IgY [10 µg/ml in 0.5% (w/v) BSA-PBS]. Rabbit anti-chicken IgY HRPO conjugate [1:5 000 in 0.5% (w/v) BSA-PBS] and 4-chloro-1-naphthol·H_2O_2 were used as the detection system.

2.4 Discussion

All active caspases, when overexpressed in heterologous expression systems, are able to undergo autoprocessing via *cis* activation (Roy *et al.*, 2014). For the *in vitro* study of caspase autoprocessing, it has been shown that inducible bacterial expression systems are ideal. This is due to the fact that autoprocessing does not occur during the early stage of induction, but as the medium acidifies, rapid autoprocessing is observed (Roy *et al.*, 2001).

Autoprocessing is required for the activation of caspases and type II MCAs, whereas type I MCAs do not depend on autoprocessing (Fuentes-Prior and Salvesen, 2004; McLuskey *et al.*, 2012). Autoprocessing has been demonstrated in the MCAs of *Leishmania* spp. (Gonzáles *et al.*, 2007; Ambit *et al.*, 2008; Zalila *et al.*, 2011; Castanys-Muñoz *et al.*, 2012), *Plasmodium falciparum* (Meslin *et al.*, 2011), *Toxoplasma gondii* (Li *et al.*, 2016), *A. thaliana* type II (Vercammen *et al.*, 2004; Bozhkov, 2005; Watanabe and Lam, 2011; Wrzaczek *et al.*, 2015), *Lycopersicon esculentum* (Wen *et al.*, 2013) *Triticum aestivum* (Piszczek *et al.*, 2012), *T. cruzi* (Laverrière *et al.*, 2012), *Tbb*MCA2 (Moss *et al.*, 2007; McLuskey *et al.*, 2012) and *Tbb*MCA4 (Proto *et al.*, 2011).

The MCA2 from *T. b. brucei* (UniProtKB ID: Q585F4) and the MCA5 from *T. vivax* (UniProtKB ID: G0U3N4), were recombinantly expressed in *E. coli*. When taking in account the addition of the bacterial expression plasmid sequence, the expected molecular weights of *Tbb*MCA2 increased from 37.5 to 42.6 kDa and *Tvi*MCA5 from 58.4 to 77.1 kDa. Upon overexpression, autoprocessing was observed and detected with both chicken anti-*Tco*MCA5 IgY and chicken anti-*Tvi*MCA5 IgY. Protein bands corresponding to the estimated sizes of the full length *Tvi*MCA5 protein, after cleavage of the N-terminal 6xHis tag, the cleavage of the C-terminal domain (21.5 kDa) from the full length protein would produce the 38 kDa protein. Subsequent cleavage of the N-terminal domain of the 38 kDa protein would result in the 29 kDa protein. The expression of *Tbb*MCA2 resulted in proteins which were of a higher molecular weight than what was expected. Thus, it was not possible to confidently assign protein domains to the cleaved products. This anomalous migration has not been reported by Moss *et al.* (2007), but has been observed in the MCA from *Nicotiana tabacum* (Acosta-Maspons *et al.*, 2014) and *A. thaliana* (Vercammen *et al.*, 2004; Watanabe and Lam, 2005).

Upon the addition of calcium, recombinant *Tbb*MCA2 produced 30 and 6 kDa bands as a result of cleavage at Lys55 and Lys268 to release the catalytic domain from the full length protein (Moss *et al.*, 2007). The Lys55 residue is only conserved in the *Tbb*MCA2 and *Tbb*MCA4, but cleavage at Lys64 was reported for *Tbb*MCA4 (Proto *et al.*, 2011). The Lys268 residue is conserved in *Tbb*MCA2 and *Tbg*MCA2, *Tbg*MCA3 and *Tev*MCA3 (Appendix A3).

Moss *et al.* (2007) were the first to recombinantly express and enzymatically characterise *Tbb*MCA2. They reported that when overexpressed in *E. coli*, *Tbb*MCA2,

with both K55G and K268G mutations, was detected in the soluble expression fraction. In our hands, recombinant *Tbb*MCA2, without the mutation of the K55 and K268 residues, was expressed within insoluble inclusion bodies. From the results of the PROSO II sequence-based solubility prediction calculator, *Tbb*MCA2 had a 38.3% chance of being insoluble when overexpressed in *E. coli* (Smialowski *et al.*, 2006). Given the predictor values for the single copy MCAs demonstrated to be insoluble, which varied between 31 and 61%, *Tbb*MCA2 could be expressed either as a soluble or insoluble protein. It was demonstrated that the full length protein was expressed in the insoluble fraction but was also expressed in the soluble fraction after the removal of the N-terminal 6xHis tag.

Western blot analysis of the expression of *Tbb*MCA2 detected five bands at 82, 68.5, 36, 30, 21.5 kDa and *Tvi*MCA5 at 77, 66, 38, 29, 21.5 kDa, using chicken anti-MCA5 IgY. This cleavage pattern resulting from autoprocessing, was thought to be due to the separation of the various domains to release the catalytic domain. This is in line with what has been reported in the literature for *Lmj*MCA, *Pf*MCA, *Tbb*MCA4 and *Les*MCA1, where autoprocessing occurs at both the N- and C-termini to produce lower molecular weight products containing the enzymatically active catalytic domain (Gonzáles *et al.*, 2007; Meslin *et al.*, 2007; Proto *et al.*, 2011; Wen *et al.*, 2013).

Affinity tags aid in the purification of recombinant peptidases and rarely affect their enzymatic activity (Uhlén *et al.*, 1992). The most commonly used affinity tag is the 6xHis which has a few advantages over other tags (Graslund *et al.*, 2008). A 6xHis tagged protein allows for the simple purification using immobilised metal affinity chromatography (Block *et al.*, 2009) and does not interfere with protein crystallography (Graslund *et al.*, 2008). The protein's structural and solubility characteristics are rarely affected by the presence of 6xHis tags when compared to that of larger affinity tags such as glutathione S-transferase and maltose-binding protein (Waugh, 2005; Nallamsetty and Waugh, 2007).

Since difficulty was encountered with the detection of the 6xHis tag of *Tbb*MCA2 in the expression fractions, both the soluble and insoluble fractions were subjected to nickel affinity purification. As expected, only the solubilised fraction bound to the nickel resin as demonstrated by the presence of proteins in the eluted fractions. Purification of recombinant *Tbb*MCA2 resulted in three protein bands, at 70.2, 32.6 and 29.2 kDa. However, no over autoprocessing was evident for *Tbb*MCA2 when compared to that seen for *Tvi*MCA5.

The nickel affinity purification of solubilised *Tvi*MCA5, resulted in multiple protein bands, predominantly at 70, 68, 38, 31, 25 and 23 kDa. These correlate with the full length protein and subsequent cleavage to release the catalytic domain from the N-terminal domain and the Pro-, Gln-, Tyr-rich C-terminal domain. The presence of numerous, less prominent protein bands, may point to secondary cleavage sites at which processing could occur after initial autoprocessing. This process was termed nickel-induced over autoprocessing and will be discussed in more detail in Chapter 3. Secondary cleavage has been demonstrated in *Lmj*MCA where autoprocessing occurs at Arg268 followed by cleavage at Arg63, Arg79 and Arg135 (Zalila *et al.*, 2011). Similarly, cleavage at Lys225 in *At*MCA4 (Watanabe and Lam, 2005), *At*MCA9 (Vercammen *et al.*, 2004; Watanabe and Lam, 2005) and MCA from *Picea abies* (Bozhkov, 2005; Watanabe and Lam, 2005) activates the peptidase, prior to cleavage at Arg190 (*At*MCA4), Arg183 (*At*MCA9) and Arg188 (*Pa*MCA).

Attempts to purify the separated domains resulting from autoprocessing were unsuccessful. This has been reported previously by Vercammen *et al.* (2004) and is thought that they had either formed complexes (Vercammen *et al.*, 2004; Klemenčič *et al.*, 2015). It was suggested by Moss *et al*. (2007) that the cleaved domains were still part of the mature protein structure and was confirmed by McLuskey *et al*. (2012) where they were shown to still be covalently linked to the main protein body. Analysis by a non-reducing SDS-PAGE by Moss *et al*. (2007) indicated that the fragments were not linked by disulphide bonds.

Antibodies were raised in chickens against the purified full length recombinant *Tvi*MCA5. The antibodies were able to detect the autoprocessing products after expression and the eluted proteins after nickel affinity purification. Since the MCAs are virulence factors, and attractive drug targets, they may also be diagnostic antigens. The MCA4 was shown to be part of the *T. b. gambiense* secretome (Geiger *et al.*, 2010) and was secreted into the culture medium of *T. b. brucei* parasites (Proto *et al.*, 2011), similar to that of oligopeptidase B (Troeberg *et al.*, 1996; Morty *et al.*, 2001), a well-known virulence factor (Coetzer *et al.*, 2008; Bastos *et al.*, 2013). The chicken antibodies raised against another well-known trypanosomal virulence factor, congopain, which is the cathepsin-L-like peptidase from *T. congolense (Tco*CATL), was shown to successfully discriminate between serum from experimentally infected and non-infected cattle in an antibody detection ELISA using an inhibition format (Eyssen, 2013). Thus, the chicken antibodies raised against *Tco*MCA5 and *Tvi*MCA5 may be used in the development of such an ELISA.

In the work described in the next chapter, the characterisation of the residues which may function in autoprocessing and peptidolytic activity of the MCA5s will be investigated by mutagenesis. Together with the kinetic analysis of both native and recombinant MCA5, insight can be gained into how the MCA5s function.

CHAPTER 3

ENZYMATIC CHARACTERISATION OF NATIVE AND MUTATED RECOMBINANT METACASPASE 5 FROM *T. CONGOLENSE* AND *T. VIVAX*

3.1 Introduction

African trypanosomosis affects both humans and livestock in rural, central Africa. There have been no new additions to the currently used chemotherapies in the last 60 years (Giordani *et al.*, 2016). As a result, drug resistance has emerged for chemotherapies used for animal treatment (Geerts *et al.*, 2001; Van den Bossche and Delespaux, 2011; Giordani *et al.*, 2016). Together with the toxicity of the currently used chemotherapies for the treatment of human infective *T. b. gambiense* and *T. b. rhodesiense*, novel chemotherapies are needed (Brun *et al.*, 2010). Progress has been made towards human chemotherapies, with two compounds in clinical trials (Maser *et al.*, 2012), but none regarding animal chemotherapies (Giordani *et al.*, 2016). In an attempt to reduce drug resistance, the use of nifurtimox-eflornithine combination therapy (NECT) over eflornithine has been recommended for human infections and the use of a sanative pair for animal infections, consisting of isometamidium chloride or homidium chloride, together with diminazene aceturate (Geerts *et al.*, 2001; Holmes *et al.*, 2004; Chitanga *et al.*, 2011).

Due to the improbability of the development of a vaccine (La Greca and Magez, 2011), focus has shifted towards the molecular targets which are essential for pathogenesis (Antoine-Moussiaux *et al.*, 2009). These molecules are known as virulence factors and are attractive targets for the development of chemotherapies (Drag and Salvesen, 2010). Numerous peptidases have been identified as virulence factors as they play an essential role in the parasite and the host-parasite interactions (McKerrow *et al.*, 2006). Due to the dual life cycle alternating between the mammalian host and insect vector, the molecular target should be as unrelated as possible to those of their mammalian hosts, to minimise any adverse side effects of the chemotherapies. Thus, in the pursuit of novel chemotherapies, the function and structure of the targeted virulence factors need to be elucidated to allow for the design of potent compounds.

The MCAs are the homologue of the metazoan caspases, which are involved in apoptosis and also have non-apoptotic roles such as cytokine maturation, inflammation

and differentiation (Lamkanfi *et al.*, 2007). Due to their presence in all kingdoms except for the metazoan kingdom, and their potential involvement in similar roles to those of the caspases, the MCAs have been identified as virulence factors. Thus, their implication in cell death, cellular processes and absence in the mammalian hosts, make the MCAs attractive targets for the development of novel chemotherapies.

The single copy MCAs are different from the multicopy MCAs as they possess a Pro-, Gln-, Tyr-rich C-terminal domain which may be involved in protein-protein interactions (Kay *et al.*, 2000). The single copy MCA5 from *T. b. brucei*, together with the multicopy MCA2 and -3 have been shown to be essential to bloodstream trypomastigotes (Helms *et al.*, 2006). However, to date, there have been no reports of the molecular characterisation of any MCA5 from the animal infective *T. b. brucei, T. congolense* or *T. vivax.*

The activation of initiator caspases involves the dimerisation of inactive caspase monomers, through their N-terminal prodomains, followed by interdomain cleavage, also known as autopeptidolytic cleavage and autoprocessing (Fig. 3.1). Dimeric executioner caspases are activated by cleavage mediated by the initiator caspases, which results in intramolecular rearrangements to form an active dimer (Tait and Green, 2010).

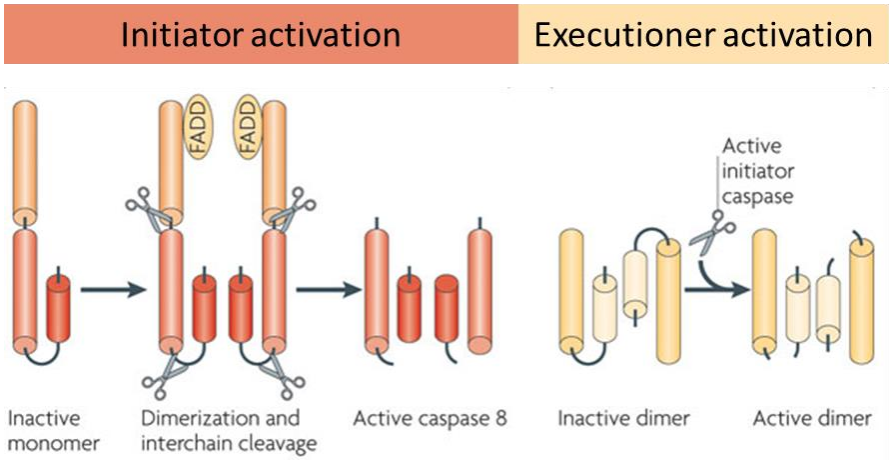


Figure 3.1: Schematic of initiator and executioner caspase activation. Dimerisation and intramolecular autoprocessing results in an active initiator caspase. Once active, the initiator caspase activates the inactive executioner dimer by peptidolytic cleavage. The fas associated death domain, FADD, is an adapter protein which is recruited to the death-inducing signalling complex (DISC) during signalling via death receptors (Gupta *et al.*, 2004). Adapted from Tait and Green (2010).

Much research has been done on plant MCAs but very little on kinetoplastid MCAs, and even fewer reported the characterisation of both recombinant and native enzymes. The residues at which the MCA is autoprocessed and the mechanism by which it

occurs are crucial for the understanding of MCA activation. Metacaspases share the conserved catalytic dyad with the caspases, but due to conserved residues within close proximity to the dyad, the question is which residues are required for the subsequent peptidolysis after activation by autoprocessing. In addition, calcium has been shown to play an important role in the autoprocessing ability and/or peptidolytic activity of the MCAs.

In the present study, the catalytic dyad residues, as well as an adjacent Cys to the catalytic Cys residue, were mutated to determine their role, if any, in autoprocessing and peptidolytic activity of the recombinant MCA5s from *T. congolense* and *T. vivax*. Using immunoaffinity purification, native *Tco*MCA5 was isolated and together with recombinant *Tco*MCA5, the substrate specificity, pH profile, inhibitor sensitivity and effect of various divalent cations was determined.

3.2 Materials and methods

3.2.1 Materials

Molecular biology: As per Section 2.2.1. The Q5® Site-Directed Mutagenesis Kit was purchased from New England Biolabs (Ipswich, MA).

***E. coli* cells:** NEB 5-alpha competent cells supplied with the Q5® Site-Directed Mutagenesis Kit.

Purification and quantification of recombinant and native proteins: As per Section 2.2.1. Digitonin was purchased from Merck (Darmstadt, Germany), the ECL western blotting substrate from Pierce (Rockford, IL, USA), and the peptidase inhibitor cocktail from Amresco (Solon, OH, USA).

Antibodies: Chicken anti-*Tco*MCA5 IgY and chicken anti-*Tvi*MCA5 IgY were produced in Section 2.2.6. The chicken anti-6xHis IgY antibody was purchased Merck (Germany) and the rabbit anti-chicken IgG HRPO conjugate from Sigma (St. Louis, MO. USA).

Parasite purification and culture: Eagle's minimal essential medium (EMEM), Gln and Pro were purchased from Sigma (St. Louis, MO, USA), foetal bovine serum (FBS) was from Gibco (Paisley, UK), DE-52 from Whatman International Ltd (UK), filters (0.2 µm) from PALL Corp (Ann Arbor, USA) and sterile disposable cell culture plasticware from Corning (NY, USA).

Peptide substrates and inhibitors: The peptide substrates benzyloxycarbonyl (Z)-Gly-Gly-Arg-7-amino-4-methylcoumarin (AMC), Z-Gly-Pro-Arg-AMC, H-Val-Gly-Arg-AMC, Z-Pro-Arg-AMC, Z-Arg-AMC, Z-Arg-Arg-AMC, t-butyloxycarbonyl (Boc)-Val-Leu-Lys-AMC, H-Ala-Phe-Lys-AMC, the inhibitors benzamidine, antipain, bestatin, N-tosyl-L-lysyl chloromethylketone (TLCK), L-*trans*-epoxysuccinyl-leucylamido(4-guanidino)butane (E-64), soyabean trypsin inhibitor (SBTI), ethylenediaminetetraacetic acid (EDTA), 4-(2-Aminoethyl)benzenesulfonyl fluoride (AEBSF), leupeptin, 1,10-phenanthroline, calpain I, iodoacetate; and porcine gelatin were purchased from Sigma (Munich, Germany). The Z-Phe-Ala-fluoromethylketone (FMK) and Z-Val-Ala-Asp-FMK inhibitors were purchased from BD Biosciences (San Jose, CA) and the black FluoroNunc™ 96-well plates from Nunc Intermed (Roskilde, Denmark).

3.2.2 Mutagenesis of the active site residues of *Tco*MCA5 and *Tvi*MCA5

It was thought that the mutation of the catalytic dyad residues might impede the ability of MCAs to autoprocess. Other researchers who have generated catalytic dyad mutants of recombinant MCAs in *Trypanosoma* spp. and *Leishmania* spp., have generally used a His→Ala and a Cys→Gly mutation (Szallies *et al.*, 2002; Gonzáles *et al.*, 2007; Lee *et al.*, 2007; McLuskey *et al.*, 2012). Primers which would facilitate the generation of the point mutants of the catalytic dyad residues (His147 and Cys202) as well as the Cys adjacent to the catalytic Cys (Cys201) were designed using the Agilent QuikChange primer design software programme (http://www.genomics.agilent.com/primerDesign Program.jsp) to achieve a Tm ≥78 °C, and compared to those designed by the above-mentioned authors. The mutagenic primers were first annealed to the gene in the pET expression plasmid, and then using the *Pfu* DNA polymerase, the primers were extended to copy the entire the plasmid. The plasmid DNA, which is isolated from dam^+ *E. coli* cells, is methylated and can be digested with DpnI. Thus, only the synthetic plasmid, which contains the desired mutation, can be taken up into the competent *E. coli* cells following transformation (Fig. 3.2).

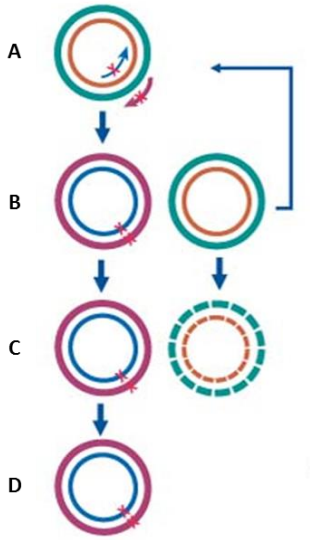


Figure 3.2: Overview of the site-directed mutagenesis process. (A) Annealing of mutagenic primers (purple and blue) encoding for site mutant to the parental plasmid template (orange and green), **(B)** followed by the amplification a new template plasmid from the parental expression plasmid (purple and blue). **(C)** DpnI digests the parental hemimethylated template, without the introduced mutation, **(D)** after which the plasmid with the mutation is transformed into competent *E. coli*. The image was obtained from the Agilent QuikChange Site-Directed Mutagenesis Kit manual.

The plasmid DNA of previously cloned *TcoMCA5* (M7pET28a) together with *TviMCA5* (V4pET32a) was used at the template for the PCR amplification with the primers outlined in Appendix B1. Briefly, the final concentrations of the PCR master mix were: 0.5 µM of each mutagenic primer, 1x Q5 high-fidelity master mix in a total reaction volume of 20 µl. The amplification was performed with incubation at 98 °C for 30 s as the initial DNA denaturation step, followed by 25 cycles of 98 °C for 10 s, 55 °C for 30 s and 72 °C for 4 min. A final elongation step at 72 °C for 2 min was subsequently carried out. A sample of the amplified plasmid which contained the gene with the introduced point mutation (2 µl) was electrophoresed on a 1% (w/v) agarose gel, containing ethidium bromide (0.5 µg/ml), in 1xTAE buffer. Thereafter, the plasmid with the point mutation (1 µl) was treated with Kinase, Ligase & DpnI enzyme mixture (KLD) for 5 min at RT. The de-methylated plasmid was transformed into competent NEB 5-alpha *E. coli* cells as described in Section 2.2.3. The plasmid DNA of five resulting colonies, from each mutant, was isolated using the GeneJet™ Plasmid Miniprep Kit, as per the manufacturer's instructions and sequenced at the Central Analytical Facilities, Stellenbosch University, South Africa, using the forward gene primer detailed in Table 2.1.

3.2.3 Expression, nickel- and immuno-affinity purification of the *Tco*MCA5 and *Tvi*MCA5 WT and mutants

Clones which were positive for the desired point mutation were expressed as outlined in Section 2.2.4 and analysed by 12.5% reducing SDS-PAGE (Laemmli, 1970) with Coomassie Blue R-250 staining and by western blot probed with chicken anti-*Tco*MCA5 IgY and chicken anti-*Tvi*MCA5 IgY (10 µg/ml).

The soluble protein lysate was subjected to immunoaffinity purification as per Section 2.2.7.2. To facilitate the concentration of eluted proteins, instead of collecting 10 sequential eluted fractions, elution buffer required for two fractions was added to the resin and agitated, prior to elution. Elution buffer, 0.1 M glycine, pH 2.8 (1.8 ml), was added to the resin followed by agitation using an end-over-end rotator for 5 min at RT. Thereafter, two 900 µl fractions were collected in microfuge tubes each containing 1 M Tris-HCl buffer, pH 8.5 (100 µl). One final fraction was collected by the addition of elution buffer (900 µl) which was allowed to flow through the resin and collected in a microfuge tube containing 1 M Tris-HCl buffer, pH 8.5 (100 µl).

The proteins from the insoluble inclusion bodies were solubilised and purified as per Section 2.2.5. Samples of each purification were electrophoresed on 12.5% reducing SDS-PAGE gels (Laemmli, 1970) and stained with Coomassie Blue R-250, with some requiring counter-staining with silver (Blum *et al.*, 1987).

3.2.4 *Trypanosoma congolense* propagation, purification and cell culture

Before the commencement of any animal procedures, ethical clearance was obtained from the University of KwaZulu-Natal animal research ethics committee (053/15/Animal). Six week old mice were kept in cages, along with enrichment material and access to food and water *ad libitum*, at the animal house at UKZN, Pietermaritzburg campus.

Female Balb/c mice were injected intraperitoneally with approximately 5 000 *T. congolense* IL3 000 bloodstream trypomastigotes with daily monitoring of parasitaemia by microscopic inspection of tail vein blood. Once parasitaemia reached 70% of the field of view, mice were sacrificed by cervical dislocation and the blood collected, by cardiac puncture, into heparin containing tubes. An equal volume of 20% (v/v) glycerol-phosphate saline glucose (PSG: 57 mM Na_2HPO_4, 3 mM NaH_2PO_4,

42 mM NaCl, 50 mM glucose, 1 mM hypoxanthine, pH 7.8) was added to the blood before storage at -80 °C or parasite purification by ion exchange chromatography.

Sterile DE-52 resin (20 ml) was washed with PBS (40 ml) followed by PSG (40 ml), in a sterile syringe containing a piece of filter paper at the bottom. Parasites were added to the resin and allowed to settle for 30 min at RT. Parasites were eluted in five fractions of 10 ml each using PSG. The parasites were pelleted by centrifugation (3 500 rpm, 15 min, 4 °C), resuspended to approximately $1x10^6$ parasites/ml in 20% (v/v) glycerol-PSG and stored at -80 °C for 7 days before being transferred to liquid nitrogen for long-term storage.

The *in vitro* culture of *T. congolense* bloodstream trypomastigotes as developed by Hirumi and Hirumi (1991) and reported by Eyford *et al.* (2011) was followed. Briefly, $1x10^5$ *T. congolense* bloodstream trypomastigotes per ml were seeded into Eagle's minimum essential medium [EMEM base powder, 2 mM glutamine, 10 mM proline, 20% (v/v) heat inactivated foetal bovine serum, 10 ml] and cultured at 37 °C and 5% (v/v) CO_2, with the medium being changed every two days.

3.2.5 Isolation of native *Tco*MCA5

Parasite fractionation in the presence of digitonin was performed as per Zalila *et al*. (2011) with some modifications. Briefly, $2x10^8$ parasites, from cell culture, were washed twice in PBS before washing in wash buffer (20 mM Tris-HCl buffer, pH 7.9, 20 mM glucose, 150 mM NaCl). The supernatant (SN) was retained, and the pelleted parasites resuspend in 100 µM digitonin/SoTe (600 mM sorbitol, 20 mM Tris-HCl buffer, pH 7.5, 2 mM EDTA, containing 1x peptidase inhibitor cocktail, 750 µl). After incubation for 5 min on ice, the total protein content from the parasite were present in the resulting SN. The total protein content was used for native immunoaffinity purification as per Section 3.2.5.1.

To the total protein content (500 µl), 0.3 M sucrose (150 µl) was added to maintain the cellular structures during centrifugation (13 000 *g*, 5 min, 4 °C). The pellet was resuspended in 500 µM digitonin/SoTE (500 µl) to obtain the proteins from the mitochondrion which were present in the supernatant after centrifugation (13 000 *g*, 5 min, 4 °C). The resulting pellet contained both the nuclear and insoluble proteins.

The proteins secreted into the culture medium were precipitated using 1 volume of 0.5 µg/ml trichloroacetic acid gently mixed with 4 volumes of medium, followed by centrifugation (18 000 *g*, 5 min, 4 °C). After washing the pellet twice with ice cold

acetone, and allowed to dry, the pellet was resuspended in reducing treatment buffer and boiled for 2 min. These samples were prepared for use in Section 5.2.13.

3.2.5.1 Immunoaffinity purification of native *Tco*MCA5

Essentially, the same protocol as described in Section 2.2.7.2 was used with some adjustments. The total parasite protein sample was added in place of the recombinant lysate and incubated with agitation, using an end-over-end rotator, for 16 h at 4 °C. After collection of the unbound proteins, the resin was washed with wash buffer (50 ml) until the absorbance readings reached a baseline A_{280} reading of 0.2. To facilitate the concentration of eluted proteins, instead of collecting 10 sequential eluted fractions, elution buffer required for two fractions was added to the resin and agitated, prior to elution. Elution buffer, 0.1 M glycine, pH 2.8 (1.8 ml), was added to the resin followed by agitation using an end-over-end rotator for 5 min at RT. Thereafter, two fractions (900 µl) were collected in microfuge tubes each containing 1 M Tris-HCl buffer, pH 8.5 (100 µl). A final fraction was collected by the addition of elution buffer (900 µl) which was allowed to flow through the resin and collected in a microfuge tube containing 1 M Tris-HCl, pH 8.5 (100 µl).

3.2.6 Enzymatic characterisation of recombinant and native *Tco*MCA5

Initial activity of the purified enzyme in each eluted fraction was measured by combining 10 µl of each fraction with MCA assay buffer (50 mM Tris-HCl buffer, pH 7.2, 150 mM NaCl, 10 mM $CaCl_2$, 5 mM DTT, 90 µl) containing 10 mM Z-Gly-Gly-Arg-AMC, in a black FluoroNunc™ 96-well plate, in duplicate. After incubation at 37 °C for 10 min, the fluorescence (Ex_{360nm} and Em_{460nm}) was measured using a FLUORStar Optima Spectrophotometer from BMG Labtech (Offenburg, Germany).

The protein concentrations of the fractions, which contained active enzyme, were determined using the BCA™ protein assay kit. The optimal enzyme concentration for the hydrolysis of the fluorogenic peptide substrate was determined by varying the concentration of the enzyme (0-5 µg/ml recombinant, 0-10 µg/ml native, in 10 µl) and incubation with MCA assay buffer containing 10 mM Z-Gly-Gly-Arg-AMC (90 µl), in triplicate, and fluorescence measured as described above.

3.2.6.1 Peptide substrate specificity

Fluorogenic peptide substrates used to determine the substrate specificity of recombinant MCA from various organisms varied from commercially available, Z-Gly-Gly-Arg AMC, Z-Gly-Arg-Arg-AMC, Z-Val-Arg-Pro-Arg-AMC, Z-Arg-Arg-AMC, Z-Arg-AMC, Boc-Val-Leu-Lys-AMC, H-Val-Leu-Arg-AMC, H-Ala-Phe-Lys-AMC, to libraries of specially designed pentapeptide fluorescence resonance energy transfer substrates derived from Abz-KARSSA-Q-EDDnp (Machado *et al.*, 2013). However, the most commonly used fluorogenic peptide substrate used for kinetoplastid MCA characterisation is Z-Gly-Gly-Arg-AMC and was, hence, the substrate of choice for the present study.

By varying the concentration of different fluorogenic peptide substrates, it is possible to determine the enzyme's maximum velocity (V_{max}) and its affinity for each substrate (K_m). This is achieved by measuring initial reaction velocity (v_o) for each of the substrate concentrations [S]. Using the Briggs and Haldane revised Michaelis-Menten equation (1925) **[A]**, a direct plot of v_o versus [S] allows for the determination of V_{max} and K_m.

$$v_o = \frac{V_{max}[S]}{[S] + K_m} \qquad \textbf{[A]}$$

The active concentration of enzyme [E_0] can be determined using a well-documented active site titrant. However, to date, there has been no report of an active site titrant specific for the kinetoplastid MCAs. Other workers assume that $[E]_t$ is equal to $[E]_0$ when determining the turnover number (k_{cat}) from V_{max} using equation **[B]** (Salvesen and Nagase, 1989):

$$k_{cat} = \frac{V_{max}}{[E]_o} \qquad \textbf{[B]}$$

Briefly, recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) and native *Tco*MCA5 (7.5 µg/ml) were incubated with MCA assay buffer at 37 °C for 10 min. During that period, a dilution series of each fluorogenic peptide substrate made in MCA assay buffer, was added to the wells of a black FluoroNunc™ 96-well plate, in triplicate, for recombinant and in duplicate for the native enzyme. The fluorescence (Ex_{360nm} and Em_{460nm}) was measured every 30 s for 60 min, using a FLUORStar Optima Spectrophotometer from BMG Labtech (Offenburg, Germany). The AMC calibration curve (Appendix B2), was used to determine the velocity, and together with the substrate concentrations, the V_{max} and K_m was determined using the Hyper32© software (1991-2003, J. S. Easterby, University of Liverpool, UK).

3.2.6.2 pH profile

The pH optimum of the recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) and native *Tco*MCA5 (7.5 µg/ml) was determined using the AMT constant ionic strength buffer (100 mM sodium acetate, 10 mM MES, 200 mM Tris-HCl, 5 mM EDTA, 5 mM DTT) at pH range of 4 to 9 (Ellis and Morrison, 1982). Enzyme (10 µl) was added to the wells of a black FluoroNunc™ 96-well plate, in triplicate for recombinant and in duplicate for native, followed by the addition of the AMT buffers, containing 10 mM Z-Gly-Gly-Arg-AMC (90 µl). After incubation at 37 °C for 10 min, the fluorescence was measured every 30 s for 15 min for recombinant *Tco*MCA5^{H147AC202G} and every 30 s for 60 min for native *Tco*MCA5.

3.2.6.3 Analysis of recombinant and native *Tco*MCA5 activity by zymography

The peptidolytic activity of recombinant (0.5 µg/ml) and native (7.5 µg/ml) *Tco*MCA5 was analysed on a 1% (w/v) porcine gelatin-containing SDS-PAGE gel (Heussen and Dowdle, 1980). The catalytic domain of congopain (*Tco*CATL, 0.5 µg/ml) was used as the positive control. An equal volume of non-reducing treatment buffer [125 mM Tris-HCl buffer, pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol] was added to samples without boiling. After electrophoresis, gels were washed twice over 1 h with 2.5% (v/v) Triton X-100 (50 ml) at RT, to remove the excess SDS. The gels were subsequently incubated at 37 °C for 16 h at 50 rpm, in AMT buffer, pH 7, 8 and 9. The gels were stained with amido black staining solution [0.1% (w/v) amido black, 30% (v/v) methanol, 10% (v/v) acetic acid] at RT for 1 h and destained with destain solution [30% (v/v) methanol, 10% (v/v) acetic acid].

3.2.6.4 Inhibitor screening

Taking into account that only a few inhibitors have been shown to effectively inhibit the MCAs from different organisms, recombinant *Tco*MCA5^{H147AC202G} was incubated with a panel of serine (TLCK, antipain, leupeptin, AEBSF, SBTI, benzamidine), cysteine (antipain, leupeptin, E-64, iodoacetic acid, calpain I), aminopeptidase (bestatin), aspartic (pepstatin), metallo (1-10 phenanthroline, EDTA), and caspase specific (Z-Phe-Ala-FMK, Z-Val-Ala-Asp-FMK) reversible and irreversible inhibitors at 1, 10 and 100 µM.

Recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) was incubated at 37 °C for 10 min with MCA assay buffer containing 10 mM Z-Gly-Gly-Arg-AMC. During that period, 1, 10 and 100 µM of the panel of reversible and irreversible inhibitors, prepared in MCA assay buffer, were added, in triplicate, to the wells of a black FluoroNunc™ 96-well plate. Thereafter, the enzyme prepared in MCA assay buffer, was added to the wells and the fluorescence measured as before (every 30 s for 60 min). The AMC standard curve (Appendix B2) and the change in velocity were used to determine the resultant inhibition of the enzyme, which was reported as the percentage inhibition compared to that of the uninhibited enzyme.

3.2.6.5 Reversible inhibition of *Tco*MCA5

Only the reversible inhibitors, antipain and leupeptin, were able to partially inhibit the activity of recombinant *Tco*MCA5^{H147AC202G}. In order to determine the inhibition kinetic constant, K_i, the recombinant enzyme was incubated with inhibitors over a range of concentrations [I], which was in 20-fold molar excess over the enzyme concentration (0-100 µM). The change in steady state velocity without inhibitor (v_0) and with the addition of inhibitor (v_i) together with equation **[C]** (Salvesen and Nagase, 1989), a plot of v_o / v_i – 1 against [I], would result in the determination of $K_{i\,(app)}$ from the gradient.

$$\frac{v_o}{v_i} = 1 + \frac{[I]}{K_{i\,(app)}} \qquad \textbf{[C]}$$

Substitution of $K_{i\,(app)}$ into equation **[D]** (Salvesen and Nagase, 1989), allows for the determination of K_i.

$$K_i = \frac{K_{i\,(app)}}{1} + \frac{[S]}{K_m} \qquad \textbf{[D]}$$

To determine the extent of reversible inhibition by antipain and leupeptin on the peptidolytic activity of *Tco*MCA5, recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) was pre-incubated at 37 °C for 10 min with MCA assay buffer containing 10 mM Z-Gly-Gly-Arg-AMC. During that period, a series of concentrations of antipain and leupeptin, made in MCA assay buffer, was added to the wells of a black FluoroNunc™ 96-well plate, in triplicate. Thereafter, the enzyme prepared in MCA assay buffer, was added to the wells and the fluorescence measured as before (every 30 s for 60 min). The AMC standard curve (Appendix B2) and the change in velocity were used to determine the resultant inhibition of the enzyme, which was reported as the percentage inhibition compared to that of the uninhibited enzyme.

3.2.6.6 Effect of divalent cations

The influence of metal ions on the activity of an enzyme can provide insight into its catalytic mechanism, as metals can serve as structural regulators, electron donors, electron acceptors or as Lewis acids (Riordan, 1977). Some MCAs have an absolute requirement for calcium for activation while others do not.

To determine the effect of divalent cations on the activity of *Tco*MCA5, recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) was pre-incubated at 37 °C for 10 min with MCA assay buffer without calcium, containing 10 mM Z-Gly-Gly-Arg-AMC. During that period, 1, 10 and 100 µM of a variety of divalent cations, prepared in MCA assay buffer, was added to the wells of a black FluoroNunc™ 96-well plate, in triplicate. Thereafter, the enzyme prepared in MCA assay buffer, was added to the wells and the fluorescence measured as before (every 30 s for 60 min). The AMC standard curve (Appendix B2) and the change in velocity were used to determine the resultant inhibition of the enzyme, which was reported as the percentage inhibition compared to that of the uninhibited enzyme.

3.3 Results

3.3.1 Mutagenesis of the active site residues of *Tco*MCA5 and *Tvi*MCA5

The generation of the catalytic dyad residue mutations, H147A and C202G, along with the Cys adjacent to the catalytic Cys, C201G, was performed in order to determine if the active site residues played a role in autoprocessing as described in Chapter 2. In addition, the double mutant of the catalytic dyad residues, H147AC202G, was generated. The engineering of *Tco*MCA5 and *Tvi*MCA5 mutants was achieved using the Q5® Site-Directed Mutagenesis Kit. The high-fidelity PCR amplification of the desired mutations are given in Appendix B3. The sequences of the generated mutants are provided in Appendix B4.

3.3.2 Expression of recombinant WT and active site mutants

A western blot of the expression of recombinant WT, using chicken anti-6xHis IgY antibodies, detected a His-tagged protein at 68 kDa for both *Tbb*MCA2 and *Tco*MCA5, and at approximately 90 kDa for *Tvi*MCA5 (Fig. 3.3, panel A) in the insoluble expression fraction. Prominent protein bands at 68 and 60 kDa for *Tco*MCA5 and

*Tvi*MCA5, as well numerous lower molecular weight protein bands, were detected using their respective chicken IgY antibodies. From literature, autoprocessing results in the cleavage of the N- and C-terminal domains to release the catalytic domain from the full length protein. Thus, the 68 and 60 kDa bands, in the insoluble expression fraction, are suggested to be the full length protein with and without the N-terminal 6xHis tag, respectively, corresponding to the proteins detected by the anti-6xHis IgY antibody. The 38 kDa protein is thought to be as a result of the removal of the C-terminal domain from the 60 kDa protein. After the removal of the C-terminal domain, 23 kDa protein, from the 38 kDa protein, the 31 kDa catalytic domain is produced.

A comparison of the soluble and insoluble expression fractions of the recombinant WT and mutants (panels B and C), revealed that apart from *Tco*MCA5^{C201G} having the same full length, 70 kDa protein band in the insoluble expression fraction, similar to that of the WT, and lacking the protein bands at 31 and 38 kDa in the soluble fraction, the mutants all had a similar protein profile to that of the WT.

No protein bands were detected after probing a western blot of the expression of BL21 (DE3) *E. coli* cells, with and without the pET-28a and pET-32a plasmids, using the chicken anti-6xHis IgY, chicken anti-*Tco*MCA5 IgY and chicken anti-*Tvi*MCA5 IgY antibodies (result not shown).

3.3.3 Affinity purification of recombinant WT and MCA5 mutants from the soluble and insoluble expression fractions

Following expression, and the detection of autoprocessed MCA in both the soluble and insoluble expression fractions, the purification strategy implemented as described in Chapter 2 was repeated. The soluble MCA was purified using immunoaffinity purification (Fig. 3.4), and the insoluble MCA was solubilised, refolded and purified using nickel affinity chromatography (Fig 3.5).

The purification of soluble MCA resulted in two prominent lower molecular weight protein bands at approximately 41 and 35 kDa, which is thought to correspond to the release of the catalytic domain from the full length protein by the cleavage of the N- and C-terminal domain, was present for all clones. Due to the very faint Coomassie stained bands present after purification of *Tco*MCA5^{C202G}, *Tvi*MCA5^{C201G} and *Tvi*MCA5^{C202G}, the SDS-PAGE gels were counter-stained with silver. The full length protein was only evident for *Tco*MCA5^{C202G} at approximately 68 kDa, and faintly for *Tvi*MCA5^{C201G} and *Tvi*MCA5^{C202G} after silver staining.

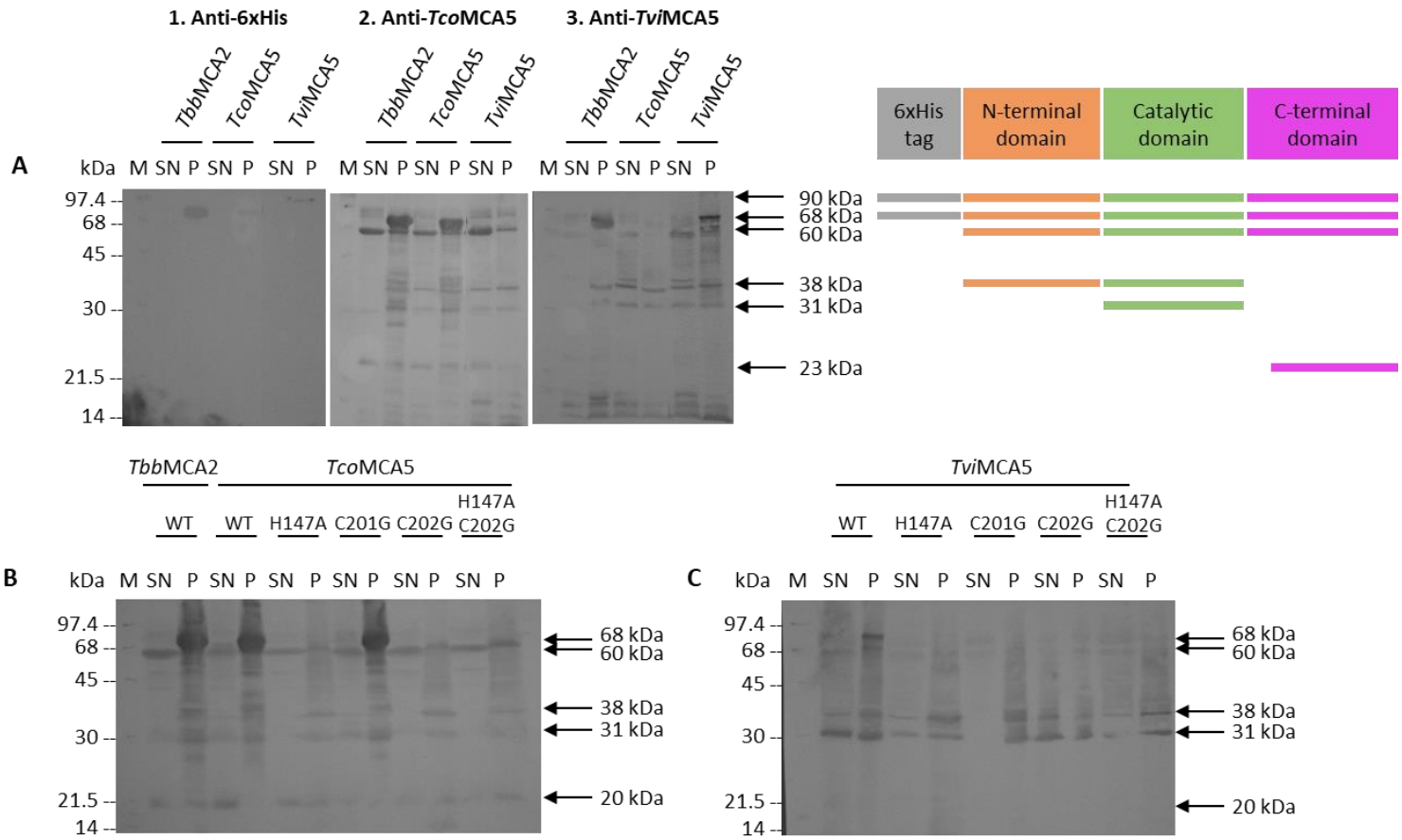


Figure 3.3: Western blots of the expression of WT MCA5 and catalytic mutants. Samples from the soluble (SN) and insoluble (P) fractions of the expression of *Tbb*MCA2, *Tco*MCA5, *Tvi*MCA5 WT along with the H147A, C201G, C202G and H147AC202G mutants of *Tco*MCA5 and *Tvi*MCA5 were electrophoresed on reducing 12.5% SDS-PAGE gels. After transfer onto nitrocellulose, blocking with 5% (w/v) milk-TBS, incubation with **(A1)** chicken anti-6xHis IgY [1:5 000 in 0.5% (w/v) BSA-PBS], **(A2 and B)** chicken anti-*Tco*MCA5 IgY and **(A3 and C)** chicken anti-*Tvi*MCA5 IgY [10 µg/ml in 0.5% (w/v) BSA-PBS] was carried out. Rabbit anti-chicken IgY HRPO conjugate [1:5 000 in 0.5% (w/v) BSA-PBS] and 4-chloro-1-naphthol·H_2O_2 were used as the detection system.

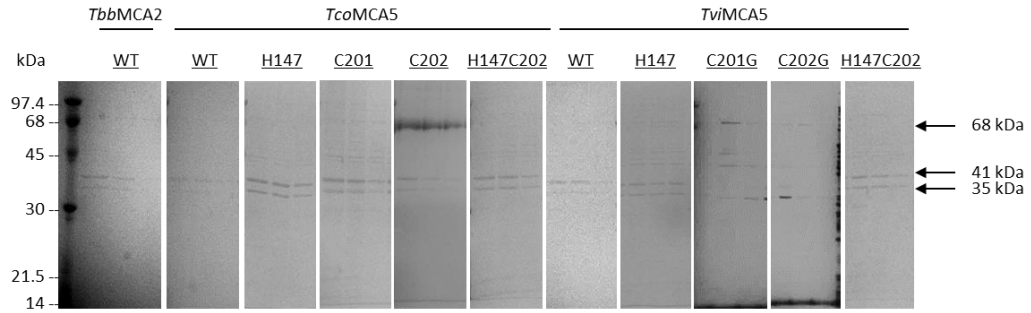


Figure 3.4: Immunoaffinity purification of the soluble WT and mutant MCA5s after recombinant expression. Samples of the three eluted fractions from the purification of soluble *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5 WT along with the *Tco*MCA5 and *Tvi*MCA5 mutants, were electrophoresed on 12.5% reducing SDS-PAGE gels and stained with Coomassie Blue R-250. The *Tvi*MCA5^{C201G} and *Tvi*MCA5^{C202G} gels were subsequently counter-stained with silver.

Similar to the observations described in Chapter 2, solubilisation, refolding and purification using nickel affinity chromatography, nickel-induced over autoprocessing was observed. That is, more protein bands were eluted than were detected in the western blot of expression samples (Fig. 3.3) and by the immunoaffinity purification (Fig. 3.4). In most clones, the 70, 68, 35, 25 and 20 kDa proteins were detected (Fig. 3.5), proteins which were visible in the western blot of expression samples. However, in the second eluted fraction, an increase in the number of eluted proteins was evident, as was shown in Chapter 2. In the case of *Tco*MCA5^{C201G} and *Tco*MCA5^{H147AC202G} and *Tvi*MCA5^{H147AC202G} clones, very little over autoprocessing was evident in the second eluted fraction, with prominent protein bands at 68, 35 and 28 kDa, corresponding to the full length and release of the catalytic domain. In fractions 4 to 10, almost pure full length protein was eluted. The elution pattern of *Tco*MCA5^{H147A} and *Tvi*MCA5 WT (panel C) differed from the other clones, with the elution of 30, 25, 21 and 14 kDa proteins in fractions 3 to 10.

3.3.4 Isolation of native *Tco*MCA5

A search of the NCBI database indicates that the MCA5 from both *T. congolense* and *T. vivax* are putative peptidases. Detection of native *Tbb*MCA5 and *Tvi*MCA5 in bloodstream trypomastigote lysate was achieved using both chicken anti-*Tco*MCA5 IgY and chicken anti-*Tvi*MCA5 IgY. The existence of *Tco*MCA5 (Eyssen, 2013) and *Tbb*MCA5 (Helms *et al.*, 2006) were demonstrated previously. However, the characterisation of *Tbb*MCA5 was not conducted. Native *Tvi*MCA5 was detected at approximately 50 kDa by chicken anti-*Tco*MCA5 IgY and at 68 and 30 kDa by chicken anti-*Tvi*MCA5 IgY, but more intensely by the latter (Fig. 3.6, panels A and B, respectively). However, native *Tbb*MCA5, at approximately 70 kDa, was only detected by chicken anti-*Tco*MCA5 IgY. As we do not currently have access to viable *T. vivax* parasites, and since *Tbb*MCA5 is not a priority enzyme in this study, only the isolation of native *Tco*MCA5 and its enzymatic characterisation was pursued.

Following the application of the total parasite protein to the chicken anti-*Tco*MCA5 IgY immunoaffinity resin, 70 and 46 kDa proteins bound to the resin and were eluted by decreasing the pH (Fig. 3.7). There was a small amount of the 46 kDa protein in the eluted fractions when compared to the amount of protein loaded onto the resin. The sizes of these proteins correspond well with the western blot of *T. congolense* bloodstream trypomastigote lysate using chicken anti-*Tco*MCA5 IgY (Eyssen, 2013). However, not all of the full length protein bound the resin as some was detected in the unbound fraction.

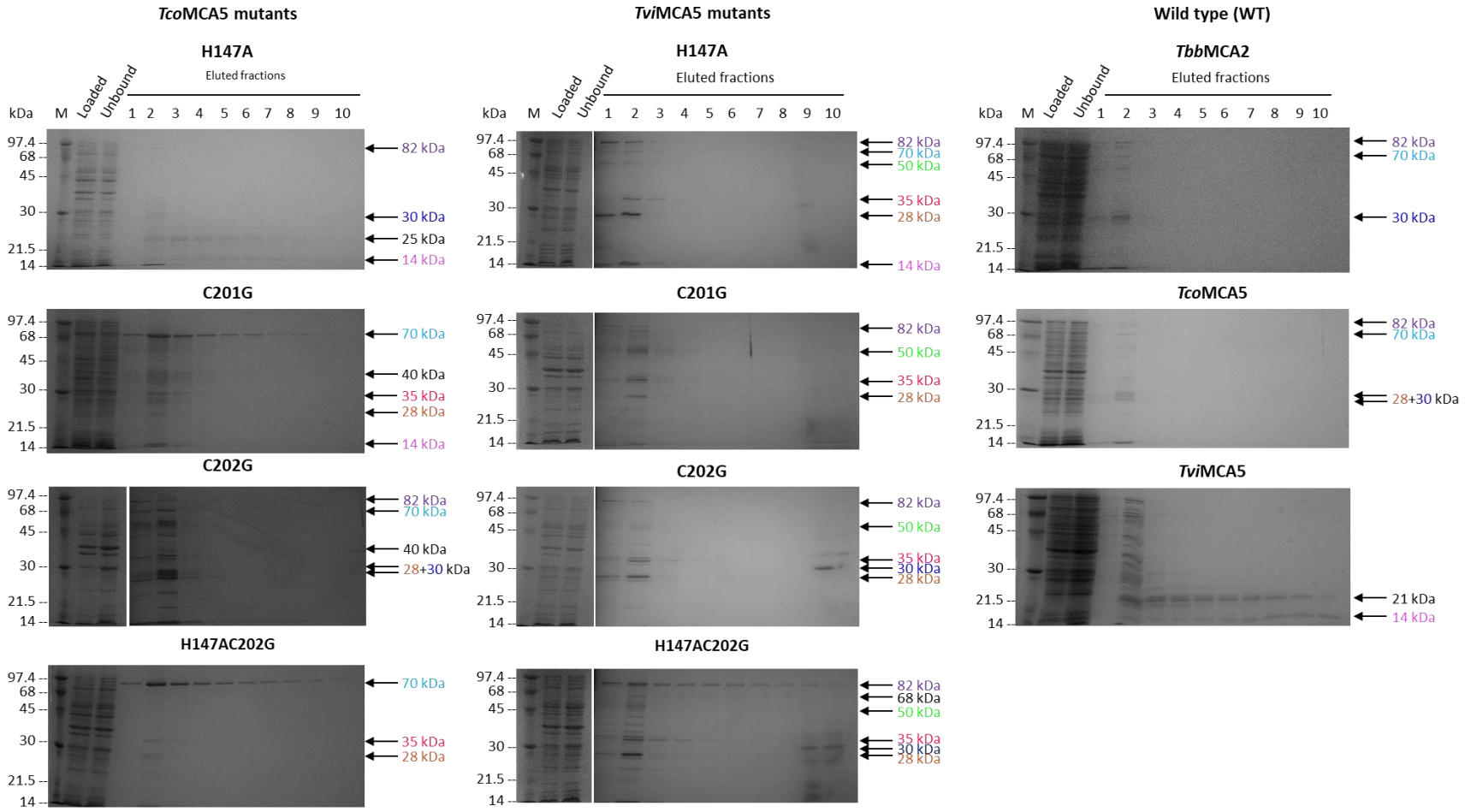


Figure 3.5: Nickel affinity purification of solubilised recombinant *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5 WT, along with *Tco*MCA5 and *Tvi*MCA5 catalytic mutants. Samples of the eluted fractions from the nickel affinity purification of recombinant MCA WT and catalytic dyad mutants were electrophoresed on 12.5% reducing SDS-PAGE gels and stained with Coomassie blue R-250. The eluted fractions of the *Tco*MCA5^{C202G} and the *Tvi*MCA5 mutants were counted stained with silver.

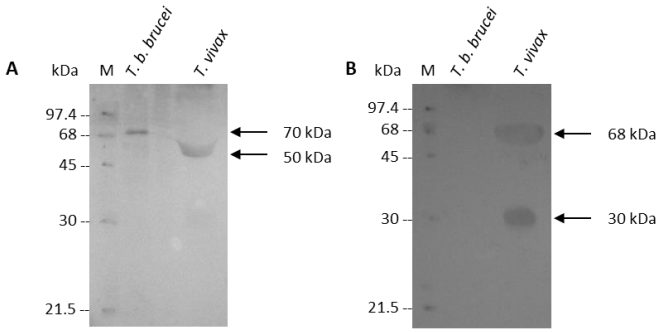


Figure 3.6: Detection of native MCA5 in *T. b. brucei* and *T. vivax* bloodstream trypomastigote lysates. Parasite lysates ($1x10^9$ parasites/ml) were electrophoresed on a 12.5% reducing SDS-PAGE gel and transferred onto nitrocellulose. After blocking with 5% (w/v) milk-TBS, the nitrocellulose was incubated with **(A)** chicken anti-*Tco*MCA5 IgY and **(B)** chicken anti-*Tvi*MCA5 IgY [10 µg/ml in 0.5% (w/v) BSA-PBS]. Rabbit anti-chicken IgY HRPO conjugate [1:5 000 in 0.5% (w/v) BSA-PBS] and 4-chloro-1-naphthol·H_2O_2 were used as the detection system.

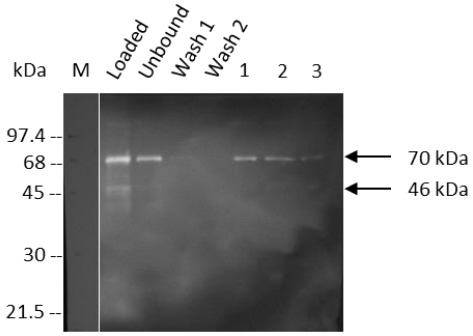


Figure 3.7: Immunoaffinity purification of native *Tco*MCA5. Samples of the total protein content loaded onto the resin, unbound fraction, wash fractions as well as the eluted fractions were electrophoresed on a 12.5% reducing SDS-PAGE gel and transferred onto nitrocellulose. After blocking with 5% (w/v) milk-TBS, the nitrocellulose was incubated with chicken anti-*Tco*MCA5 IgY [10 µg/ml in 0.5% (w/v) BSA-PBS]. Rabbit anti-chicken IgY HRPO conjugate [1:5 000 in 0.5% (w/v) BSA-PBS] and ECL were used as the detection system.

3.3.5 Enzymatic characterisation

Both the WT and mutants were enzymatically characterised based on their P_1 site specificity, pH profile as well as which inhibitors they were sensitive to.

Hydrolysis of the Z-Gly-Gly-Arg-AMC substrate was only detected in the second fraction of the double mutant *Tco*MCA5^{H147AC202G} purified using nickel affinity chromatography. In addition, hydrolysis was detected with the immunoaffinity purified, native *Tco*MCA5. To prove that these two enzymes were indeed the only active MCAs, the second eluted fraction from each of the 'inactive' clones, were tested alongside those of the 'active' MCAs. It was found that 0.5 µg/ml *Tco*MCA5^{H147AC202G} (pink) and 7.5 µg/ml native *Tco*MCA5 (green) were active with a discernible increase in velocity

(Fig. 3.8, panels A and B, respectively), whilst the eluted fractions from the 'inactive' clones showed no activity. Despite the higher enzyme concentration, native *Tco*MCA5 had a much slower velocity, 3.51 µM/s, compared to recombinant *Tco*MCA5^{H147AC202G}, at 145 µM/s.

The fact that the clone in which both the catalytic residues were mutated, was shown to be the most active, supports the suggestion of the existence of a secondary catalytic Cys (Vercammen *et al.*, 2004; Proto *et al.*, 2011; Piszczek *et al.*, 2012). The Cys residue in question, Cys81 in *Tco*MCA5, is in close proximity to the catalytic Cys202 and adjacent Cys201 (Appendix B5), and is conserved in the MCAs found in *Trypanosoma* spp. and *Leishmania* spp. (Appendix A3). Further support for this suggestion is that neither the catalytic His nor Cys were shown to be essential for the 'trypsin-like' activity of native *Ldn*MCA (Lee *et al.*, 2007).

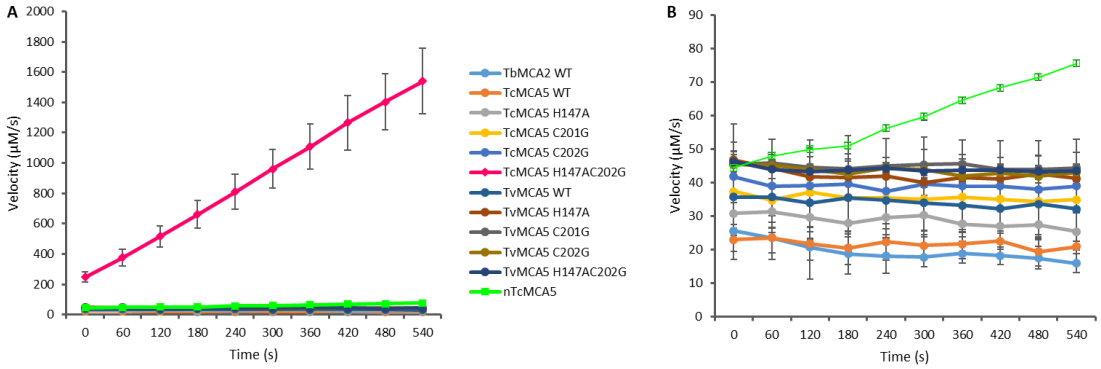


Figure 3.8: Enzymatic activity of the purified recombinant WT, mutants and native *Tco*MCA5. From the purifications using nickel affinity chromatography of: recombinant WT and all mutants of both *Tco*MCA5 and *Tvi*MCA5, together with the immunoaffinity purified native *Tco*MCA5, the second eluted fraction (0.5 µg/ml and 7.5 µg/ml for recombinant and native, in triplicate and duplicate, respectively) was incubated with MCA assay buffer, containing 10 µM Z-Gly-Gly-Arg-AMC and the resultant fluorescence measured (Ex_{360nm} and Em_{460nm}). The equation of the **(A: pink)** recombinant *Tco*MCA5^{H147AC202G} trendline is given by y = 145.88x with a correlation coefficient of 0.9997 and for **(B: green)** native *Tco*MCA5, y = 3.5158x with a correlation coefficient of 0.9868.

3.3.5.1 Substrate specificity

The preference of recombinant *Tco*MCA5^{H147AC202G} and native *Tco*MCA5 for different amino acid residues at the P_1, P_2 and P_3 sites was investigated. Due to the low activities obtained for Lys in the P_1 site, only the activities with Arg in the P_1 site are reported in Table 3.1. Despite the higher affinity for both Z-Pro-Arg-AMC and Z-Arg-Arg-AMC, by recombinant *Tco*MCA5^{H147AC202G}, Z-Gly-Gly-Arg-AMC was characterised by a higher turnover number. However, recombinant *Tco*MCA5^{H147AC202G} cleaved Z-Arg-Arg-AMC two times more efficiently than Z-Gly-Gly-Arg-AMC. The most commonly reported

fluorogenic peptide substrate used for kinetoplastid MCAs is Z-Gly-Gly-Arg-AMC, which was selected for the remaining enzymatic characterisations.

Since a high concentration of native *Tco*MCA5 was required to observe any activity, it was only tested against Z-Gly-Gly-Arg-AMC, and in duplicate. As seen in the velocities reported in Fig. 3.8, native *Tco*MCA5 hydrolysed the fluorogenic peptide substrate very slowly (3.917 µM/s), with a low turnover of 0.52 s^{-1}, as well as having a 14 times lower substrate affinity compared to that of the recombinant *Tco*MCA5^{H147AC202G}. The catalytic efficiency constant, $k_{cat}/K_{m,}$ is useful in comparing the relative rates of an enzyme acting on different substrates, but is not to be used to compare two different enzymes acting on the same substrate (Eisenthal *et al.*, 2007). Thus, the low value for native *Tco*MCA5, compared to recombinant *Tco*MCA5^{H147AC202G}, does not necessarily mean that Z-Gly-Gly-Arg-AMC is a poor choice of substrate. The Lineweaver-Burke plots are provided in Appendix B6.

Table 3.1: Substrate specificity of recombinant *Tco*MCA5^{H147AC202G} and purified native *Tco*MCA5. Recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) and native *Tco*MCA5 (7.5 µg/ml) were incubated in MCA assay buffer (50 mM Tris-HCl buffer, pH 7.2, 150 mM NaCl, 10 mM $CaCl_2$, 5 mM DTT) before the addition of 0 to 30 µM of various fluorogenic peptide substrates in triplicate and duplicate, respectively. The resultant fluorescence from the hydrolysis of the AMC fluorophore, over 60 min, was measured at Ex_{360nm} and Em_{460nm}. The kinetic parameters, K_m and V_{max} were determined using the Hyper32® software.

Substrate (µM)	*Tco*MCA5^{H147AC202G}			Native *Tco*MCA5		
	K_m (µM)	k_{cat} (1/s)	k_{cat}/K_m (1/s.µM)	K_m (µM)	k_{cat} (1/s)	k_{cat}/K_m (1/s.µM)
Z-Gly-Gly-Arg-AMC	120	871	7.26	8.4	0.52	0.062
H-Val-Gly-Arg-AMC	72.9	420	5.76	ND	ND	ND
Z-Gly-Pro-Arg-AMC	171	808	4.73	ND	ND	ND
Z-Pro-Arg-AMC	24.0	211	8.79	ND	ND	ND
Z-Arg-Arg-AMC	13.2	175	13.3	ND	ND	ND
Z-Arg-AMC	134	408	3.04	ND	ND	ND

ND: not determined due to limited quantity of native *Tco*MCA5 enzyme available.
The Boc-Val-Leu-Lys-AMC and H-Ala-Phe-Lys-AMC substrates had K_m and V_{max} values of less than 0.02 µM, and were omitted from the table.

3.3.5.2 pH profile

Using the Z-Gly-Gly-Arg-AMC substrate, along with the constant ionic strength AMT buffers, the pH profile of recombinant *Tco*MCA5^{H147AC202} was determined, using triplicate samples, in the absence and presence of calcium. Based on the results of the recombinant enzyme and limited amount of native enzyme, the pH profile of native *Tco*MCA5 was determined at six pH values instead of eleven, using duplicate samples, without the addition of calcium.

Recombinant *Tbb*MCA2 has a pH optimum of 7.7 (Machado *et al.*, 2013; Gilio *et al.*, 2017) due to the two ionising groups of the catalytic dyad with pK_{e1} = 6.7 (His) and pK_{e2} = 8.7 (Cys) (Machado *et al.*, 2013). The addition (panel A) and omission of calcium (panel B) slightly affected the activity of recombinant *Tco*MCA5^{H147AC202G}, but was shown to be most active between pH 7 and 9. Moderate activity was found between pH 6 and 7, with very little activity below pH 5.5 (Fig. 3.9). The omission of calcium (panel B) resulted in slightly increased velocities when compared to those obtained with the addition of calcium (panel A). As such, due to the limited amount of native *Tco*MCA5, the pH profile was determined without the addition of calcium. The pH profile for native *Tco*MCA5 (panel C) was very similar to that of recombinant *Tco*MCA5^{H147AC202G}, with a broad pH optimum between 6.5 and 9, at a lower velocity, even after an extended period of measurements. This was expected following the low velocities observed in the range of 50 to 70 µM/s in Fig. 3.8, panel B.

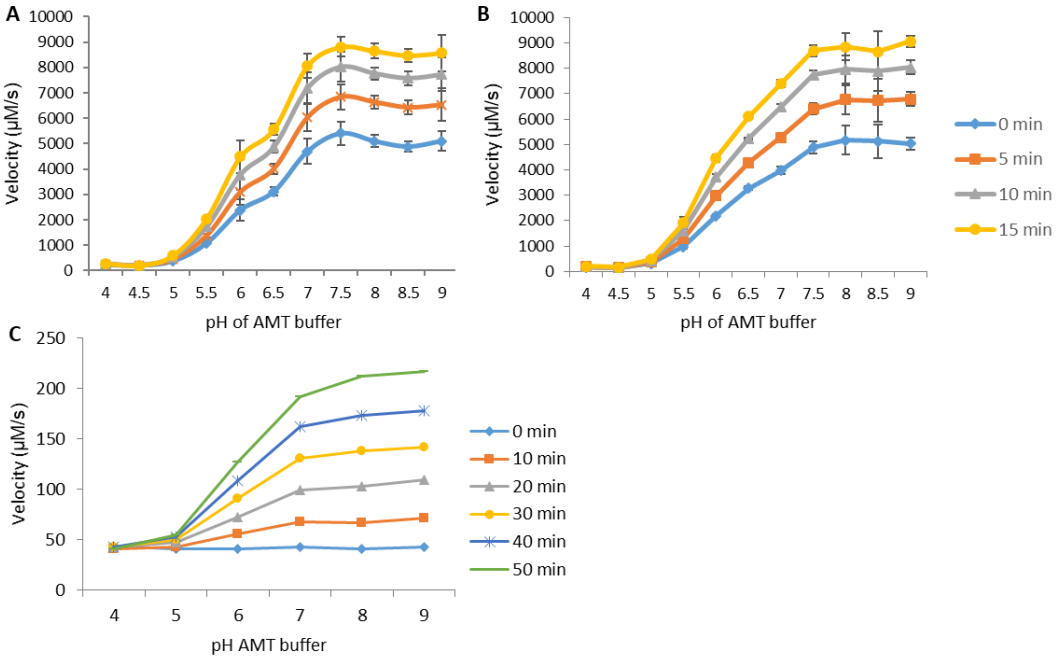


Figure 3.9: The pH profile of recombinant *Tco*MCA5^{H147AC202G} and native *Tco*MCA5. Recombinant *Tco*MCA5^{H147AC202G} (**A and B;** 0.5 µg/ml, in triplicate) and native *Tco*MCA5 (**C;** 7.5 µg/ml, in duplicate) were incubated in constant ionic strength AMT buffers (100 mM sodium acetate, 10 mM MES, 200 mM Tris-HCl, 5 mM EDTA, 5 mM DTT) varying in pH from 4 to 9 and containing Z-Gly-Gly-Arg-AMC (10 µM). In **(A),** the AMT buffer was supplemented with 10 mM $CaCl_2$ whereas **(B and C)** was without. The resultant fluorescence from the hydrolysis of the AMC fluorophore by recombinant and native *Tco*MCA5 over 15 and 60 min, respectively, was measured at Ex_{360nm} and Em_{460nm}.

The estimated p*I* of the full length *Tbb*MCA2 is 5.8 compared to 8.4 for *Tco*MCA5. A probable cause for the broad pH profile, and the deviation from the bell-shaped curve of *Tbb*MCA2 may be due to the difference in amino acid composition of the N-terminal

domain and the longer C-terminal domain of MCA5 (Fig. 2.7 and Appendix A3) which remains attached to the main protein body after autoprocessing (McLuskey *et al.*, 2012). The p*I*-values of these domains differ between *Tbb*MCA2 (N-terminal domain, 6; C-terminal domain, 5), *Tco*MCA5 (N-terminal domain, 9.8; C-terminal domain, 8) and *Tvi*MCA5 (N-terminal domain, 10.2; C-terminal domain, 9).

3.3.5.3 Gelatin digestion

In order to determine which of the proteins, purified using nickel- and immuno-affinity chromatography for recombinant *Tco*MCA5^{H147AC202} and immuno-affinity chromatography for native *Tco*MCA5, were active, zymography was employed. Activity is visualised by the digestion of the gelatin substrate, shown as clear areas in the amido black stained gelatin-containing non-reducing SDS-PAGE gel. As seen in Fig. 3.9, both recombinant *Tco*MCA5^{H147AC202G} and native *Tco*MCA5 were active between pH 7 to 9, the three zymograms were therefore incubated at pH 7, 8, and 9. The recombinant catalytic domain (CD) of congopain, *Tco*CATL, which served as the positive control, showed activity at approximately 22 kDa. Activity of the MCA5 was observed at approximately 97, 63, 42 and 38 kDa with no discernible difference in activity observed between the three buffers (Fig. 3.10).

Due to the addition of gelatin to the SDS-PAGE gels, the molecular weights are only approximations. Referring to Appendix B7, the two active protein bands of native *Tco*MCA5 correspond well to the 68 and 46 kDa protein bands detected after its immunoaffinity purification as detected using chicken anti-*Tco*MCA5 IgY (Fig. 3.7). Recombinant *Tco*MCA5^{H147AC202G} was active as two protein bands which corresponds well to the 32 and 28 kDa protein bands seen in Fig. 3.5. Samples from the nickel- and immuno-affinity purifications of both recombinant WT and clones of the other mutants, were also electrophoresed on gelatin containing SDS-PAGE gels, with gelatin digestion only evident for the recombinant *Tco*CATL CD control (results not shown).

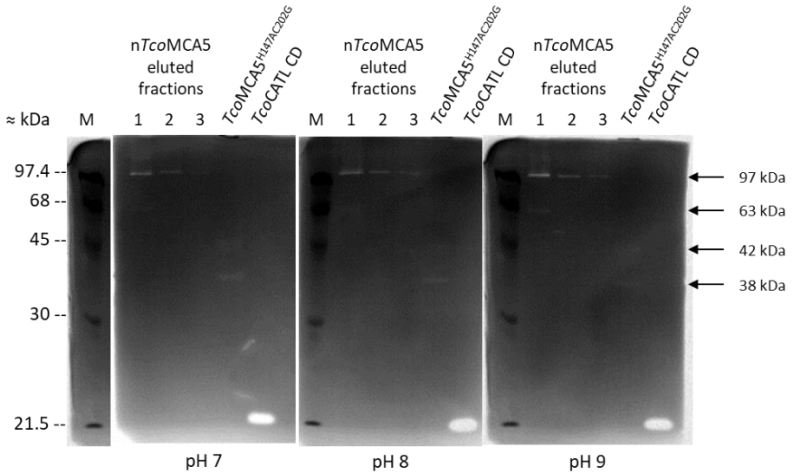


Figure 3.10: Gelatin containing SDS-PAGE to visualise the peptidolytic activity of native *Tco*MCA5 and recombinant *Tco*MCA5^{H147AC202G}. A sample from each of the eluted fractions of purified native *Tco*MCA5, second eluted fraction of nickel purified recombinant *Tco*MCA5^{H147AC202G}, together with recombinant *Tco*CATL CD (0.5 µg/ml), was electrophoresed on a non-reducing 12.5% SDS-PAGE containing 1% (w/v) porcine gelatin. After washing twice with 2.5% (w/v) Triton X-100 and incubation with AMT buffers pH 7, 8 and 9, the gels were stained with Amido Black.

3.3.5.4 Inhibitors

A panel of both reversible and irreversible inhibitors where incubated with recombinant *Tco*MCA5^{H147AC202G} to determine which would have an effect on enzyme activity. As evident from Fig. 3.11, only the serine and cysteine specific inhibitors, leupeptin and antipain were able to reduce the enzyme activity of recombinant *Tco*MCA5^{H147AC202G} by more than 20% at 10 µM. An increase in inhibitor concentration to 100 µM, did not increase the extent of inhibition. Both the general caspase inhibitor, Z-Val-Ala-Asp-FMK, and negative control for caspase inhibitors, Z-Phe-Ala-FMK, showed no inhibition. Even the classical cysteine specific inhibitor, E-64 was unable to reduce activity (Appendix B8). The velocities and standard deviations of each inhibitor are given in Appendix B9.

The inhibition constant, K_i, for antipain was 119.1 µM and 72.99 µM for leupeptin (Appendix B10), indicating that leupeptin was a better inhibitor than antipain.

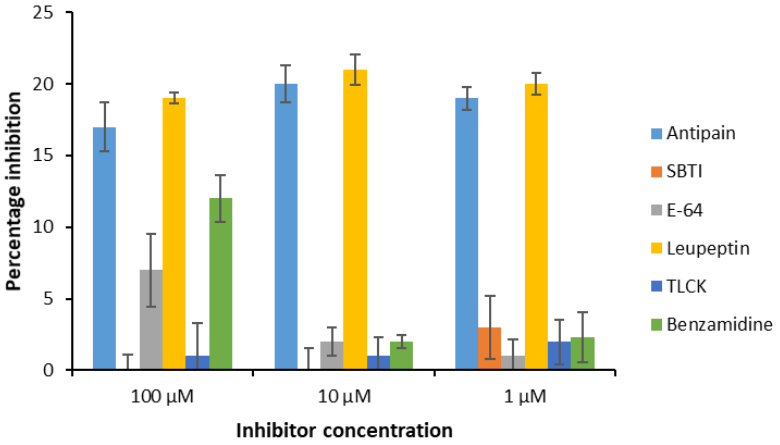


Figure 3.11: Sensitivity of recombinant *Tco*MCA5^{H147AC202G} to a variety of peptidase inhibitors. Recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) was incubated in MCA assay buffer containing 10 µM Z-Gly-Gly-Arg-AMC, before the addition of 1, 10 and 100 µM of various peptidase inhibitors, in triplicate, diluted in assay buffer without substrate. The resultant fluorescence from the hydrolysis of the AMC fluorophore, over 60 min, was measured at Ex_{360nm} and Em_{460nm}. No inhibition was measured for calpain I, bestatin, AEBSF, 1,10 phenanthroline, Z-Phe-Ala-FMK, Z-Val-Ala-Asp-FMK and iodoacetic acid. Percentage inhibition was determined using 0 mM inhibitor as 0% inhibition.

3.3.5.5 Effect of divalent cations

Some MCAs depend on calcium activation for peptidolytic activity (Bozhkov, 2005; Moss *et al.*, 2007; He *et al.*, 2008; Ojha *et al.*, 2010; Meslin *et al.*, 2011; Proto *et al.*, 2011; Laverrière *et al.*, 2012; Wen *et al.*, 2013) whilst others do not (Vercammen *et al.*, 2004; Gonzáles *et al.*, 2007; Klemenčič *et al.*, 2015).

In the absence of calcium, recombinant *Tco*MCA5^{H147AC202G} was active, however, the addition of 1 and 5 mM calcium, decreased its activity by 10%, but at 10 and 25 mM 100% activity was observed (Fig. 3.12, panel A). Increasing the calcium concentration beyond 25 mM resulted in a steep drop in activity, where at 100 mM, half its activity had been lost. The decrease in velocity is given in Appendix B11. As such, the pH profile of native *Tco*MCA5 was tested without the addition of calcium in the MCA assay buffer.

In an attempt to understand why extensive autoprocessing is evident in the second eluted fraction after nickel affinity purification of the solubilised recombinant MCA5s, the effect of other divalent cations was investigated. Nearly complete inhibition of Z-Gly-Gly-Arg-AMC hydrolysis, was observed for 1 mM Cd^{2+} and Zn^{2+}, 58% for Cu^{2+}, 54% for Fe^{2+} and 32% for Ni^{2+} when added to recombinant *Tco*MCA5^{H147AC202G} (Fig. 3.12, panel B). When increased to 10 mM, all cations, with the exception of Mg^{2+}, Mn^{2+} and Ca^{2+}, decreased the activity of *Tco*MCA5^{H147AC202G} by more than 90%. The

decrease in velocity is given in Appendix B12. This trend might be due to the size of the ions where Cd>Zn>Cu>Ni>Fe>Mn>Ca>Mg and the resultant percentage inhibition Cd=Zn=Cu>Fe> Ni>Mg=Mn=Ca are almost directly related.

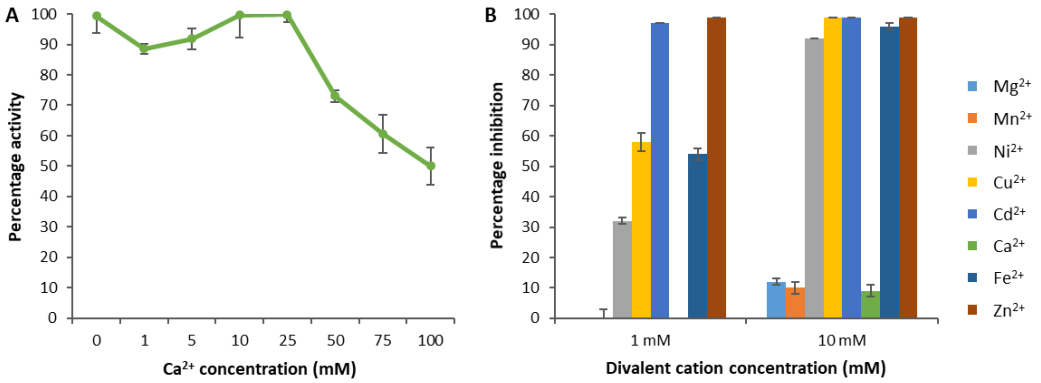


Figure 3.12: Effect of divalent cations on the enzymatic activity of recombinant *Tco*MCA5^{H147AC202G}. Recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) was incubated in MCA assay buffer containing 10 µM Z-Gly-Gly-Arg-AMC, before the addition of **(A)** 0 to 100 mM $CaCl_2$ or **(B)** 1 and 10 mM of various divalent cations, in triplicate, diluted in assay buffer without substrate. The resultant fluorescence from the hydrolysis of the AMC fluorophore, over 60 min, was measured at Ex_{360nm} and Em_{460nm}. Percentage activity is determined using 0 mM $CaCl_2$ as 100% active and percentage inhibition using 0 mM inhibitor as 0% inhibition.

3.4 Discussion

Due to the structural similarities to the caspases and their absence in the metazoan kingdom, it was hypothesised that the MCAs may function in a similar manner, and have hence been described as virulence factors (Mottram *et al.*, 2003). The MCAs' involvement in both apoptotic (Proto *et al.*, 2011; Zalila *et al.*, 2011; Laverrière *et al.*, 2012) and non-apoptotic (Helms *et al.*, 2006; Ambit *et al.*, 2008; Lee *et al.*, 2010) events have been demonstrated. Thus, targeting the MCA for the development of novel chemotherapies, may cause a disruption in the signalling pathways of the parasite, and be ideal for parasite control. As such, it is imperative to characterise the MCAs' activation mechanisms and kinetic characteristics.

Autoprocessing of MCAs generally functions to remove the N- and C-terminal domains from the catalytic domain. However, it has been suggested that the cleavage of the N-terminal domain is independent of the peptidase's autoprocessing ability (Gonzáles *et al.*, 2007). The MCAs do not necessarily require autoprocessing, nor dimerisation for activation of their peptidolytic activity (Moss *et al.*, 2007), and are not all activated by calcium (Vercammen *et al.*, 2004; Gonzáles *et al.*, 2007). It has been demonstrated that recombinant *Tbb*MCA2 undergoes autoprocessing, but this is not necessary for activity, whereas autoprocessing is key for activation of caspases (Moss *et al.*, 2007).

No *Tbb*MCA2 autoprocessing has been demonstrated *in vitro*, but it cannot be ruled out as its occurrence may be restricted to specific points of the cell cycle or under specific *in vivo* induction conditions (Moss *et al.*, 2007; Ambit *et al.*, 2008). It is a possibility that *in vivo*, autoprocessing plays a role in MCA targeting and/or protein interactions and thus, function in an indirect regulatory manner.

The catalytic domain of *Lmj*MCA, when expressed in yeast, was shown to be more active than the full length protein (Ambit *et al.*, 2008). This is in line with the suggestion that processing is required for optimal enzymatic activity (Gonzáles *et al.*, 2007). A similar result was obtained for *Pf*MCA1 when expressed in *E. coli* (Meslin *et al.*, 2007). Mutagenesis of the sites of autoprocessing in *Tbb*MCA2 resulted in an active enzyme but not as active as when autoprocessing, to release the catalytic domain, had occurred (Gilio *et al.*, 2017). This, however, was not the case for *Tcr*MCA3 and -5 where autoprocessing was not observed but the enzymes were still active (Kosec *et al.*, 2006). Mutagenesis studies have been conducted to determine if the catalytic residues play a role in autoprocessing. The individual mutations of the catalytic His of the single yeast metacaspase, YCA1, and the residues of the catalytic dyad of *Lmj*MCA, did not impair their ability to undergo autoprocessing (Gonzáles *et al.*, 2007; Wong *et al.*, 2012). In addition the Cys adjacent to the catalytic Cys, Cys201, of *Lmj*MCA was shown not to be essential for peptidolysis but did have an influence on the catalytic Cys which was not discussed further (Gonzáles *et al.*, 2007). The two Cys residues, Cys201 and Cys202, of *Tco*MCA5 and *Tvi*MCA5 do not interact directly with each other. Instead Cys201 and Cys202 form π-sulfur interactions with His200 and Tyr141, respectively, and is illustrated in Appendix B13. To date, there has been no record of a recombinant MCA clone in which both the catalytic His and Cys were mutated.

Considering the conserved sequences between the MCAs of *Leishmania* spp., together with the preference of Arg at the P_1 site of substrates, the domain assignment of the multiple bands seen in the expression samples of the recombinant WT and mutants of *Tco*MCA5 and *Tvi*MCA5 were based on that assigned for *Lmj*MCA autoprocessing. Firstly, C-terminal cleavage at Arg298, followed by N-terminal cleavage at Arg63, Arg79 and Arg135, is required for the release of the active catalytic domain (Zalila *et al.*, 2011). Upon the overexpression of *Tco*MCA5 and *Tvi*MCA5 in *E. coli,* several bands were detected, which included the full length protein, with and without the N-terminal 6xHis tag at 68 and 60 kDa, respectively. Through autoprocessing, the removal of the 23 kDa C-terminal domain from the 60 kDa protein

would produce the 38 kDa protein comprising the catalytic domain and the N-terminal domain. Removal of the N-terminal domain would result in the 31 kDa catalytic domain which is close to the expected size of *Tco*MCA5 and *Tvi*MCA5. To confirm the domains assigned to these proteins, and determine the sites of autoprocessing, N-terminal sequencing needs to be performed in future studies.

No significant differences in the autoprocessing patterns were observed upon expression of the WT and the catalytic residue mutants. Significant differences were only observed in the elution profiles of the purifications of the soluble and insoluble expression fractions. In the case of the immunoaffinity purification of the soluble expression fraction, only *Tco*MCA5^{C202G} possessed the full length protein. The nickel affinity purification of both the WT and mutant MCAs, from the solubilised expression fraction, resulted in nickel-induced over autoprocessing, as reported in Chapter 2, for the WT and each mutant with the exception of *Tco*MCA5^{H147AC202G}. The nickel affinity purification of the *Tco*MCA5 in which both the catalytic dyad residues had been mutated, resulted in no nickel-induced over processing, but rather, only the full length 68 and 35, 28 kDa proteins were eluted. The 35 kDa protein corresponds to the cleavage of the N-terminal 6xHis tag as well as the C-terminal domain. Removal of the N-terminal domain from this protein would result in the 28 kDa protein.

Following the isolation of total protein content from *T. congolense* bloodstream trypomastigotes, native *Tco*MCA5 was immunoaffinity purified as a full length 68 kDa and a 45 kDa protein. The removal of the C-terminal domain from the full length protein is thought to produce the 45 kDa protein.

When analysed on a zymogram, both the 68 and 45 kDa protein bands of native *Tco*MCA5, and the 35 and 28 kDa protein bands of *Tco*MCA5^{H147AC202G} showed gelatin digestion. None of the other purified MCA WT and mutant clones showed any gelatin digestion, which correlates with the result of only the recombinant *Tco*MCA5^{H147AC202G} and native *Tco*MCA5 showing any activity towards the Z-Gly-Gly-Arg-AMC substrate. The native *Tco*MCA5 had a low reaction velocity and affinity for the fluorogenic peptide substrate, which may have been caused by the small amount of active catalytic domain as well as the active full length protein which was visualised in the zymogram. To date, this is the first report of MCA, or caspase, digestion of gelatin.

The existence of a secondary catalytic Cys has been suggested due to its close proximity to the catalytic Cys, and its conservation in all kinetoplastid species (Szallies *et al.*, 2002; Belenghi *et al.*, 2007). It is though that the secondary catalytic Cys may

be required for autoprocessing or for its peptidolytic activity (Szallies *et al.*, 2002). The function of this secondary catalytic Cys has been shown to rescue *At*MCA9 activity despite *S*-nitrosylation of the catalytic Cys when endogenous levels of nitric oxide are high (Belenghi *et al.*, 2007). The fact that the only active recombinant MCA was the double catalytic mutant, provides more evidence towards the existence of a secondary catalytic Cys which supports that report that the 'trypsin-like' activity of *Ldn*MCA is not completely dependent on the catalytic dyad residues (Lee *et al.*, 2007). In the type II MCA from *Triticum aestivum*, *Tae*MCAII, it was shown, through mutagenesis studies of the catalytic Cys140 and secondary catalytic Cys23, that autoprocessing was dependent on catalytic Cys140 and that the secondary catalytic Cys23 was essential for catalytic activity (Piszczek *et al.*, 2012). However, this does not rule out the role played by the Cys adjacent to the catalytic Cys in peptidolysis.

To date, there is no suitable inhibitor of MCA that could be used for active site titration. However, Machado *et al.* (2013) for *Tbb*MCA2, Laverrière *et al.* (2012) for *Tcr*MCA3 and *Tcr*MCA5, as well as Vercammen *et al.* (2006) for *At*MCA9, reported k_{cat} values but failed to mention the active site titrant. Klemenčič *et al.* (2015) used the irreversible inhibitor Z-Phe-Arg-FMK as an active site titrant for the recombinant orthocaspase 1 from *Microcystis aeruginosa* PCC 7806, whilst Watanabe and Lam (2011) showed that biotin-Phe-Pro-Arg-CMK was able to abolish the ability of recombinant *At*MCA4 to cleave Z-Gly-Arg-Arg-AMC. However, they did not use this to determine the k_{cat}. Thus, until it is shown that either Z-Phe-Arg-FMK or biotin-Phe-Pro-Arg-CMK are able to effectively inhibit *Tco*MCA5, we will assume that $[E]_t$ is equal to $[E]_0$.

In line with literature on *At*MCA4 (Vercammen *et al.*, 2004), *At*MCA9 (Watanabe and Lam, 2005) and *Tbb*MCA2 (Machado *et al.*, 2013), it was demonstrated that recombinant *Tco*MCA5^{H147AC202G} had a preference for Arg over Lys at the P_1 site. The tripeptide fluorogenic substrates had greater reaction velocities than their dipeptide counterparts. The Z-Arg-AMC had the second highest substrate affinity, but hydrolysis occurred at a low velocity. At the P_2 site, Pro was favoured over Gly, and at the P_3 site, Gly was preferred over Val. These findings are in line with what was reported for recombinant *Tbb*MCA2 where Val, Ala, Glu, Leu and Pro are preferred at the P_3 site and Gly, Leu, Phe, Pro, Arg and Lys at the P_2 site (Machado *et al.*, 2013).

Despite the lower affinity of recombinant *Tco*MCA5^{H147AC202G} for Z-Gly-Gly-Arg-AMC, (119.6 µM) compared to 25.7 µM reported for recombinant *Tbb*MCA2, hydrolysis occurred at a lower velocity of 435.6 µM/s for *Tco*MCA5^{H147AC202G} compared to 3372.5 µM/s of *Tbb*MCA2, (McLuskey *et al.*, 2012). Purified native *Tco*MCA5

possessed a higher substrate affinity of 8.384 μM for and slower hydrolysis (3.917 μM/s) of Z-Gly-Gly-Arg-AMC compared to recombinant *Tco*MCA5^{H147AC202G}. This may be due to the incomplete autoprocessing to release the active catalytic domain of native *Tco*MCA5 as seen in the zymogram when compared to that of recombinant *Tco*MCA5^{H147AC202G}.

The pH optimum of MCA may give insight to its localisation, and hint at which functions it might be involved in. The different processed forms of *Lmj*MCA are targeted to different subcellular compartments (Zalila *et al.*, 2011; Castanys-Muñoz *et al.*, 2012). The optimal peptidolytic activity for MCAs has been reported to be between pH 7 and 9 (Vercammen *et al.*, 2004; Moss *et al.*, 2007; Machado *et al.*, 2013; Wen *et al.*, 2013), with the exception of *At*MCA9, which has a pH optimum of 5.5. In contrast to the bell-shaped curve demonstrated by most MCAs, native the *Ldn*MCA1 was active over a broad pH range (Lee *et al.*, 2007). This is similar to the pH profile of recombinant *Tco*MCA5^{H147AC202G} and native *Tco*MCA5 showing activity from pH 6 to 9. This board pH range may be due to the difference in p*I* values of the N-and C-terminal domains between the single- and multicopy MCAs which stay attached to the main protein body after autoprocessing (McLuskey *et al.*, 2012). The C-terminal domains of the single copy MCAs are longer than those of the multicopy MCAs and have a higher p*I*. In addition, the N-terminal domains of the single copy MCAs have a higher p*I* than those of the multicopy MCAs. It is, thus, hypothesised that the increased p*I* of both the N- and C-terminal domains may contribute to the observed activity across a broad range of pH values.

The most successful inhibitors of recombinant *Tbb*MCA2 (Moss *et al.*, 2007), *Tbb*MCA4^{S219C} (Proto *et al.*, 2011) and native *Ldn*MCA (Lee *et al.*, 2007) were found to be leupeptin, antipain, TLCK with 70 to 90% inhibition, AEBSF and chymostatin with 30 to 45% inhibition. Only antipain, 119.1 μM, and leupeptin, 72.99 μM, had measurable inhibition kinetic constants for recombinant *Tco*MCA5^{H147AC202G}, equating to a 20% inhibition by antipain and 21% by leupeptin. Each of the other inhibitors, specific for different peptidase classes and at different concentrations, which were tested, had no effect on the activity of recombinant *Tco*MCA5^{H147AC202G}. Due to the limited supply of native *Tco*MCA5, inhibitor efficacy was not determined.

Studies have shown that a number of MCAs have an absolute calcium requirement for activity (Bozhkov, 2005; Moss *et al.*, 2007; He *et al.*, 2008; Ojha *et al.*, 2010; Meslin *et al.*, 2011; Proto *et al.*, 2011; Laverrière *et al.*, 2012; Wen *et al.*, 2013). The few MCA exceptions which do not require calcium for activity are *Lmj*MCA (Gonzáles *et al.*,

2007), *At*MCA9 (Vercammen *et al.*, 2004) and the orthocaspase, *Mae*OC1 (Klemenčič *et al.*, 2015). Calcium has been shown to induce an allosteric conformational change which functions to stabilise substrate binding (McLuskey *et al.*, 2012). As such, the calcium dependency of some MCAs is suggested to function as a mechanism for the control of peptidolysis (Moss *et al.*, 2007). Autoprocessing of recombinant *At*MCA4 produces four protein bands which lack the N-terminal affinity tag (Watanabe and Lam, 2011). After the addition of calcium to the expression sample, an increase in the number of *At*MCA4 protein bands is observed with cleavage at both the N- and C-termini (Watanabe and Lam, 2011).

The calcium-induced processing of *At*MCA4 is similar to what was seen in the second eluted fraction in the present study's nickel-affinity purification, hence the term, nickel-induced over autoprocessing. Various authors have used nickel-, and in some cases cobalt-, affinity chromatography to purify plant, yeast, leishmanial and trypanosomal recombinant MCAs (Vercammen *et al.*, 2004; Kosec *et al.*, 2006; Gonzáles *et al.*, 2007; Lee *et al.*, 2007; Moss *et al.*, 2007; Proto *et al.*, 2011; Laverrière *et al.*, 2012; McLuskey *et al.*, 2012; Piszczek *et al.*, 2012; Wong *et al.*, 2012; Wen *et al.*, 2013). However, the SDS-PAGE gels showing the elution profiles were not reported in these studies.

It was demonstrated that calcium was not required for the activity of recombinant *Tco*MCA5^{H147AC202G}. In fact, activity decreased between 6 and 7% upon the addition of calcium up to 10 mM, while between 10 and 25 mM no reduction in activity was observed. Above 25 mM, a steep decrease in activity was observed with increasing calcium concentration, while at 100 mM, activity was reduced by 50%.

In an attempt to understand why nickel-induced over autoprocessing occurred during nickel-affinity purification, the effect of different divalent cations was investigated. The addition of 1 mM Cd^{2+} and Zn^{2+} as well as 10 mM Cu^{2+}, Fe^{2+} and Ni^{2+} resulted in the almost complete inhibition (>95%) of recombinant *Tco*MCA5^{H147AC202G}. The addition of other divalent cations, Cu^{2+}, Mg^{2+}, Mn^{2+} and Zn^{2+} failed to affect the activity of recombinant *Tbb*MCA2 (Moss *et al.*, 2007). Addition of these cations to YCA1 and the MCA from *Lycopersicon esculentum*, *Les*MCA1 (Wen *et al.*, 2013), showed no effect when visualised on a reducing SDS-PAGE gel (Wong *et al.*, 2012). In the case of recombinant *At*MCA2, the addition of Mg^{2+} and Mn^{2+} increased the activity slightly as determined by fluorescence (Watanabe and Lam, 2011).

The His-select® nickel affinity resin contains approximately 20 to 40 µmol divalent nickel ions per 1 ml of resin. At 1 mM Ni^{2+}, a 30% reduction in activity was measured. Thus, the low concentration of nickel ions in the resin would be insufficient to significantly reduce the activity of the recombinant *Tco*MCA5^{H147AC202G}. Bearing in mind that the recombinant *Tco*MCA5^{H147AC202G} did not require calcium to be active, it is hypothesised that the binding of the 6xHis tagged recombinant enzyme to the nickel resin immobilises it and allows for the interaction of the MCA with itself, as reported for *Lmj*MCA by Casanova *et al*. (2015), and other MCA molecules in its immediate surroundings. This is in line with what was suggested by Watanabe and Lam (2011), that recombinant MCA may exhibit minor catalytic activity which could initiate a cascade reaction of peptidolytic processes. Thus, when two proenzyme molecules are in close proximity, one slightly active proenzyme could cleave the second molecule (Watanabe and Lam, 2011). This suggests that *Tco*MCA5 requires autoprocessing to release the active catalytic domain.

This may allow for the recombinant *Tco*MCA5 to over autoprocess at secondary cleavage sites or even non-specifically. Thus, upon elution, there is no intact active catalytic domain (35 and 28 kDa bands), left to hydrolyse the fluorogenic peptide substrate. This over autoprocessing was not evident in the immunoaffinity purification of the recombinant WT (Section 2.3.6) and, as such, must be attributed to the nickel ions in the resin.

With this information about substrate and pH preferences, the modelling of these proteins including their catalytic residues, secondary catalytic Cys residues and the identification of other residues which may play a role in peptidolysis was performed. In addition, docking of an inhibitor library, which was designed for *Tbb*MCA2 by Berg *et al.* (2012), was performed for both *Tbb*MCA5 and *Tvi*MCA5 in an attempt to identify suitable pharmacophores for the development of effective anti-MCA inhibitors as described in Chapter 4.

CHAPTER 4

DOCKING STUDIES OF MCA2 INHIBITORS WITH MCA2 FROM *T. B. BRUCEI* AND MCA5 FROM *T. CONGOLENSE* AND *T. VIVAX*

4.1 Introduction

Trypanosomal parasites' ability to evade the adaptive immune response of the host, by the process of antigenic variation, impedes the development of vaccines against both human and animal Africa trypanosomosis, HAT and AAT (Morrison *et al.*, 2009). Thus, prevention through vector control and screening for infections, followed by the appropriate treatment are the only options for the control of AT.

The currently used trypanocides have been in use for over 60 years (Barrett *et al.*, 2011), with the exception of the nifurtimox-eflornithine combination therapy for HAT (Priotto *et al.*, 2009). None of the currently used chemotherapies are effective against both *T. b. gambiense* and *T. b. rhodesiense*, as well as both the initial and final disease stages. Side effects of HAT chemotherapies vary in severity, with the arsenical based melarsoprol causing reactive encephalopathy in 5 to10% of all treated patients. It is uncertain how the HAT chemotherapies, pentamidine, suramin and melarsoprol actually kill trypanosomes (Barrett *et al.*, 2007). It has been shown previously that the HAT chemotherapies suramin (Morty *et al.*, 1998) and pentamidine (Ashall, 1990; Morty *et al.*, 1998) as well as the AAT chemotherapy, diminazene aceturate (Morty *et al.*, 1998) were able to inhibit the activity of *Tbb*OPB. However, it is probable that these drugs have several modes of action through the interaction with multiple targets within the parasite (Gilbert, 2013). Due to the indiscriminate use of AAT chemotherapies and the length of time that the same drugs have been in use, drug resistance has been reported (Giordani *et al.*, 2016). Thus, it is imperative to develop novel chemotherapies to treat and control AT (Field *et al.*, 2017).

Drug discovery is predominately based on the identification of individual targets in the infectious organism which can be affected by small molecule ligands, which in turn can be modified to achieve better affinity for the target, resulting in a more potent inhibitor with drug-like characteristics (Brun *et al.*, 2011; Lounnas *et al.*, 2013). The most important feature of a target for drug discovery is that it must be essential to the parasite, i.e. be a virulence factor. Other important features to consider is the

availability of structural information of the target, the target's ability to be inhibited by drug-like molecules and be selective for the parasite only (Frearson *et al.*, 2007; Wyatt *et al.*, 2011; Gilbert, 2013).

Various methodologies are being pursued in the discovery of novel chemotherapies (Matthews, 2015; Field *et al.*, 2017). The target-based methodology relies on the identification of a target, and along with its structural information (Sliwoski *et al.*, 2013), structure-based drug design (SBDD) can be used for the identification of binding ligands, and the subsequent optimisation of ligand structure to increase binding affinity (Skinner-Adams *et al.*, 2016). The SBDD methodology is one of the most promising methodologies to identify suitable ligands for target proteins (Bleicher *et al.*, 2003; Foloppe *et al.*, 2006; Klebe, 2006) and involves ligand docking, pharmacophore design, and ligand design methods (Sliwoski *et al.*, 2013). Together with the three dimensional (3D) structure of the target, the interaction energies can be calculated for target specific ligands by molecular docking techniques (Kalyaanamoorthy and Chen, 2011). A positive outcome of SBDD methodologies is the identification of the important residues involved in binding, which can inform modification of existing ligands to achieve a greater binding affinity for the target (Jain *et al.*, 2011; Blaney, 2012).

Knowledge of the target structure and the sites of interaction with specific ligands may assist in the deduction of the roles played by the target (Urwyler, 2011). Cellular peptidases (McKerrow et al., 2008) and multiple enzymes, involved in various biochemical pathways, have been validated as targets for the development of novel chemotherapies (Field *et al.*, 2017), with some progress being made in the identification of chemical inhibitors. The MCAs have been identified as virulence factors and are attractive targets for drug discovery due to their absence in the animal kingdom and their potential involvement in parasite cellular processes. *Trypanosoma b. brucei* possesses five MCA genes, of which *Tbb*MCA2 is well characterised. The single copy MCA5 differs from MCA2 by the presence of a Pro-, Glu-, Tyr-rich C-terminal domain, which is thought to be involved in protein-protein interactions (Kay *et al.*, 2000). The MCA sequences of the human and animal infective *Trypanosoma* spp., *T. cruzi* and *Leishmania* spp. show a high level of conservation (Fig. 2.1 and Appendix A3).

A library of ligand inhibitors was designed based on the P_1 site specificity of *Tbb*MCA2 for basic Arg and Lys residues prior to the 3D structure being solved (Berg *et al.*, 2012). The ligands were tested against the purified recombinant *Tbb*MCA2 as well as in a phenotypic assay using *in vitro* cultured *T. b. brucei, T. cruzi, L. infantum* and

P. falciparum parasites, which achieved micromolar IC_{50} values, between 2.1 and 50 µM. However, ligand optimisation was not pursued further. The 3D structure of *Tbb*MCA2 has since been solved (McLuskey *et al.*, 2012), but the binding site and affinity of the Berg ligands for the *Tbb*MCA2 target has not been established.

In the present study, the validation of ligand binding, from the previously designed Berg library, into the active site of the 3D structure of *Tbb*MCA2 (solved by X-ray diffraction (McLuskey *et al.*, 2012)) was performed using molecular docking software. This software evaluates the ligand conformation, when docked in the active site by specific scoring functions, and modifies the torsional, translational and rotational degrees of freedom of the ligands, until a minimum energy is achieved (Kapetanovic, 2008; Yuriev *et al.*, 2009; Huang and Zou, 2010). In addition, the affinity of the Berg ligands for the MCA5s from animal infective *T. congolense* and *T. vivax*, is investigated.

4.2 Materials and Methods

4.2.1 Materials

Commercial library: The structures of the inhibitors and fluorogenic peptide substrates used in literature for the characterisation of various MCAs were obtained from the PubChem compound database (Kim *et al.*, 2016) and can be found in Appendix C1.

Berg library: The library was obtained from The Binding Database, (http://www.bindingdb.org/jsp/dbsearch/PrimarySearch_pubmed.jsp?pubmed_submit=Search&pubmed=20167486) (Gilson *et al.*, 2016) and can be found in Appendix C2 The IC_{50} values reported in the analyses are of the ligands tested against the recombinant *Tbb*MCA2 as published by Berg *et al*. (2012).

***Tbb*MCA2 3D structures:** The 3D structures, solved using X-ray diffraction, of*Tbb*MCA2 with the C213G mutant with, 4AF8 (1.4 Å), and without a Sm^{3+} ion, 4AFP (2.1 Å), and the C213A mutant with, 4AFV (1.5 Å), and without a calcium soak, 4AFR (1.6 Å), were obtained from the RCSB Protein Data Bank (www.rcsb.org)(Berman *et al.*, 2000) as submitted by McLuskey *et al*. (2012).

Discovery Studio docking software: The Centre for High Performance Computing (Cape Town, South Africa) is acknowledged for the provision of the computational resources necessary to conduct this work.

4.2.2 Homology modelling

The most reliable computational method for predicting protein structure from its sequence is homology modelling (Bordoli *et al.*, 2008). This involves the identification of templates using a BLAST search, followed by the alignment and superimposing of the template to the query sequence using both structure and sequence alignment methods. The end result is a 3D model of the query protein.

The SwissModel software (Biasini *et al.*, 2014) was used to create 3D models for *Tco*MCA5 and *Tvi*MCA5. Together with multiple sequence alignment, the scoring functions, Qmean and GMQE (Global Model Quality Estimation), indicators of the reliability of the modelled protein, were reported (Benkert *et al.*, 2009; Benkert *et al.*, 2011).

The Ramachandran plots for *Tco*MCA5 and *Tvi*MCA5 were generated by Discovery Studio (Dassault Systèmes BIOVIA), and compared to those of *Tbb*MCA2 reported in the Protein Data Bank by McLuskey *et al*. (2012). The quality of the generated models was verified using the ProSA-web protein structure analysis software whereby an overall quality score for a query structure is reported (https://prosa.services.came.sbg.ac.at/prosa.php) (Sippl, 1993; Wiederstein and Sippl, 2007).

4.2.3 Identification of residues involved in catalysis

The residues which are involved in catalysis and calcium binding, in addition to the gatekeeper and latch residues, which function in the removal of the N-terminal domain from across the active site, were highlighted in Appendix A3 and listed in Table 4.1. The 3D structures, solved by X-ray diffraction, of *Tbb*MCA2, 4AF8 and 4AFP had a catalytic Cys mutation at residue 213 to Gly, and in 4AFV and 4AFR to an Ala. This residue in each of the four *Tbb*MCA2 models was changed back to Cys to best simulate the environment in the active site.

Table 4.1: Identification of residues in *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5 which are involved in catalysis.

	*Tbb*MCA2	*Tco*MCA5 and *Tvi*MCA5
Gatekeeper	Tyr31	Tyr19
Latch/S_1 binding	Ser156	Ser145
S_1 binding	Asp95, Asp211	Asp84, Asp200
Calcium binding	Asp173, Asp189, Asp190, Asp220	Asp162, Asp178, Asp179, Asp209
Secondary catalytic Cys	Cys92	Cys81
Catalytic dyad	His158	His147
	Cys213	Cys202
Adjacent Cys	Cys212	Cys201

4.2.4 Protein and ligand preparation

The *Tbb*MCA2 models, 4AF8, 4AFP, 4AFV, 4AFR, and the homology modelled *Tco*MCA5 and *Tvi*MCA5, were prepared using the 'Prepare Protein' protocol from Discovery Studio (Dassault Systèmes BIOVIA). The protocol allowed for the optimisation of side chain conformation, the retention or removal of water molecules, the modeling of any missing loops and side chain atoms, and the protonation of the structure by the prediction of the site pK_as at the optimum pH of *Tbb*MCA2, 7.7, as determined by Machado *et al*. (2013). To assist in differentiation between the different protein models, the structures were coloured differently, 4AF8 (pink), 4AFP (white), 4AFV (orange), 4AFR (blue), *Tco*MCA5 (purple) and *Tvi*MCA5 (turquoise).

The structural water molecules are usually located in deep pockets of the protein 'receptor' structure and mediate the formation of hydrogen bonds between the ligand and receptor binding site (Bissantz *et al.*, 2010). Strongly bound water molecules are often conserved across multiple crystallographic structures (Ferreira *et al.*, 2015) as is evident for *Tbb*MCA2, where three conserved water molecules line the bottom of the active site in all four 3D structures (McLuskey *et al.*, 2012). As such, each of the proteins were prepared in four ways. Firstly, with all the water molecules present; secondly, with all water molecules removed; thirdly, only the conserved water molecules present; and finally, from the prepared structure with all water molecules present, all but the conserved water molecules were removed. These preparations will be referred to as 'all water', 'no water', 'conserved water, prepared', and 'conserved water without preparation' for simplicity.

The ligand inhibitors comprising the Berg library, as well as the substrates and inhibitors comprising the commercial library, were prepared using the 'Prepare Ligands' protocol from Discovery Studio (Dassault Systèmes BIOVIA). The protocol

allows for the generation of multiple tautomers and stereoisomers for each ligand molecule, the correction of valencies and formal charges associated with each functional group. Finally, the generated 3D conformation of each inhibitor, substrate and ligand was prepared for use in docking studies.

4.2.5 Binding site identification

The binding sites were defined using the receptor cavity protocol using Discovery Studio (Dassault Systèmes BIOVIA). This method is based on an 'eraser' algorithm which functions to remove all points which are not in contact with the protein 'receptor' (Venkatachalam *et al.*, 2003). The binding site which was identified within the active site cavity, comprising the catalytic dyad, secondary and adjacent Cys residues as well as the S_1 binding residues, was selected for docking.

Upon the addition of calcium, a conformational change occurs in *Tbb*MCA2 which functions to remove the N-terminal domain from across the active site (McLuskey *et al.*, 2012). This involves the S_1 binding residue closest to the catalytic His residue, which functions as a latch, and the gatekeeper Tyr residue in the N-terminal domain (McLuskey *et al.*, 2012). Taking this into consideration, some of the N-terminal domain was removed, residues 1 to 35 for *Tbb*MCA2 and residues 1 to 25 for MCA5, the valence corrected, subjected to the 'Prepare Protein' protocol with each of the four different water preparations, followed by the re-evaluation of the binding sites. This final structure was used for all subsequent docking studies.

4.2.6 CDocker

Protein 'receptors' and ligands are flexible and may occupy multiple conformations in solution. In addition, upon ligand binding, the protein 'receptor' may undergo a conformational change. Most docking algorithms function by the docking of flexible ligands into the rigid binding site of the protein 'receptor'. Another method, which is more realistic but requires more CPU time, is flexible docking whereby various protein 'receptor' conformations are used for docking with flexible ligand molecules. The CDocker algorithm was used in this study, which, through the use of random initial ligand placement and CHARMm forcefield, docks ligand molecules into the defined binding site of the protein 'receptors'. Using the default settings in the CDocker protocol from Discovery Studio (Dassault Systèmes BIOVIA), the ligands from both the Berg and the commercial libraries where docked into the binding site of the 3D structures of

*Tbb*MCA2 4AF8, 4AFP, 4AFV and 4AFR as well as the homology modelled structure of both *Tco*MCA5 and *Tvi*MCA5.

4.2.7 Common feature pharmacophore

A pharmacophore is defined as the group of steric and electronic features which is required for the activation or inhibition of a receptor. A group of structurally diverse ligands are able to bind to a common site on a receptor should they share a common pharmacophore. Novel ligands targeting specific sites can be designed using these pharmacophore models. A common feature pharmacophore was generated using the same named protocol from Discovery Studio (Dassault Systèmes BIOVIA). This was achieved using the ligands which had the best CDocker interaction energies with corresponding low IC_{50} values and docked in a similar spatial position.

4.3 Results

4.3.1 *Tbb*MCA2 from the Protein Data Bank

To date the only 3D structures of MCAs, solved by X-ray diffraction, are YCA1 from *Saccharomyces cerevisiae*, 4F6P and 4F6O (Wong *et al.*, 2012), and *Tbb*MCA2, 4AF8, 4AFP, 4AFV and 4AFR (McLuskey *et al.*, 2012).

The *Tbb*MCA2 peptidase is dependent on calcium for activation, and in 4AF8, samarium was used to define the calcium binding site (McLuskey *et al.*, 2012). The Sm^{3+} ion coordinates four Asp residues and two water molecules (Fig. 4.1, panel A). Three structural water molecules were observed to be conserved between the four 3D structures, solved by X-ray diffraction, of *Tbb*MCA2 and were found to line the bottom of the active site pocket (Fig. 4.1, panel B). These molecules interact with the catalytic His158, the S_1 binding residue Asp211, the secondary catalytic Cys92 as well as with Gly157 and Gly159. The oxyanion hole of the caspases is formed by the backbone nitogen atoms of the catalytic Cys and the conserved Gly238, which is adjacent the catalytic His (Walker *et al.*, 1994; Fuentes-Prior and Salvesen, 2004). The catalytic His of the MCAs is adjacent to a Gly residue which is conserved in all the kinetoplastid MCAs (Appendix A3). As such, Gly159 may form part of the oxyanion hole of *Tbb*MCA2.

Residues 269 to 275 form a disordered region which corresponds to the ‘280-loop’. Due to the disorder, this loop was not solved in the crystal structure (Fig. 4.1, panel C and D). Upon soaking *Tbb*MCA2 4AFV crystals in calcium, a shift in position of the

'280-loop' was observed (Fig. 4.1, panels D and E). As a result, a 5.6 Å, 110° shift occurs at the Cα position of Gly280 caused by the change in direction at the Cα position of Ala279 (Fig. 4.1, panel F) as reported by McLuskey *et al*. (2012).

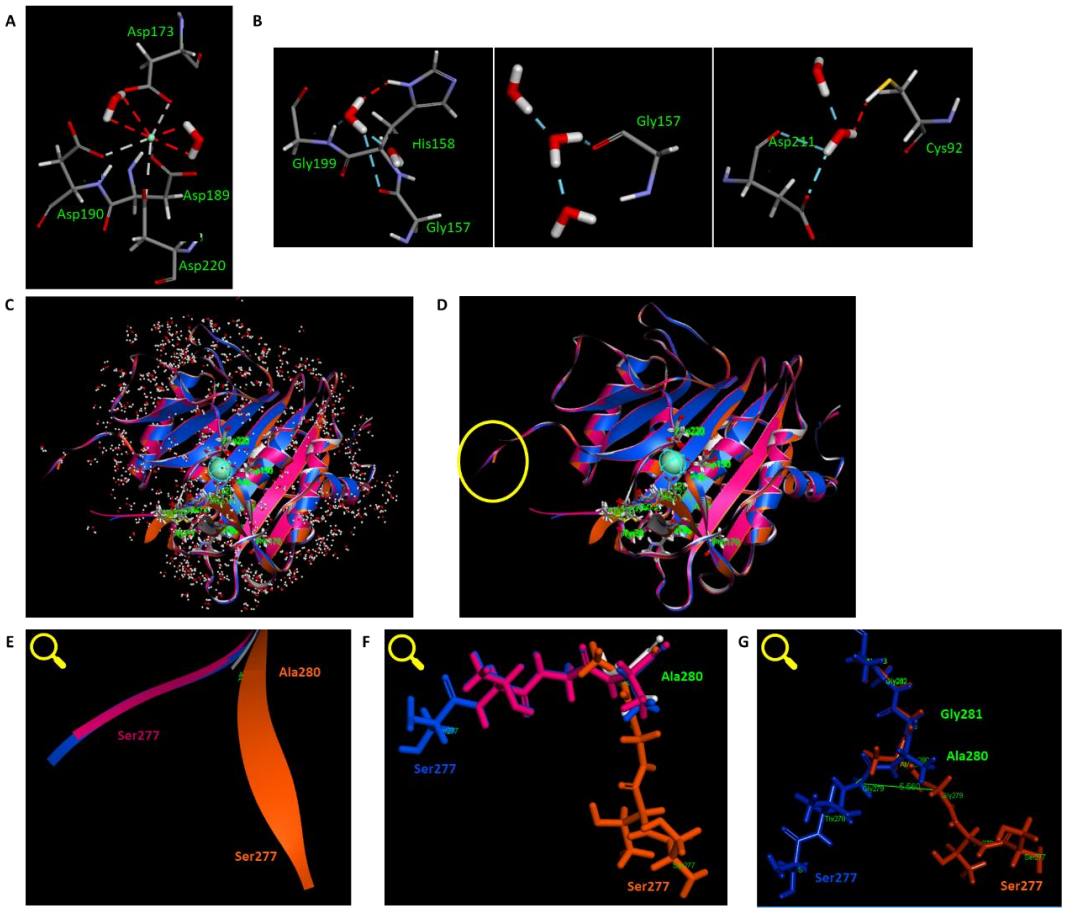


Figure 4.1: Structural features of the 3D structures of *Tbb*MCA2 for docking. (A) Water molecules and conserved Asp residues interacting with Sm^{3+} ion (teal) in *Tbb*MCA2 4AFP. **(B)** Interactions of the three conserved water molecules with the residues in the active site of *Tbb*MCA2. Overlay of the 3D structures, solved by X-ray diffraction, of *Tbb*MCA2 4AF8 (pink), 4AFP (white), 4AFV (orange) and 4AFR (blue) (McLuskey *et al.*, 2012) **(C)** with and **(D)** without all the water molecules. The C-terminal of the '280-loop' is highlighted in **(D)** and the magnified view shown in **E**, **F** and **G**. Magnification of the '280-loop' **(E)** ribbon structure, **(F)** atomic structure and **(G)** comparison change in direction of 4AFV and 4AFR due to calcium soaking. Structures were obtained from the Protein Data Bank (Berman *et al.*, 2000), visualised and analysed using Discovery Studio (Dassault Systèmes BIOVIA).

4.3.2 Homology modelling

The SwissModel alignment of *Tco*MCA5 and *Tvi*MCA5 resulted in a 48 to 49% sequence identity to the four 3D structures, solved by X-ray diffraction, of *Tbb*MCA2. The results from the CDocker analysis of the four *Tbb*MCA2 models, together with the four water preparations, with both the commercial and Berg libraries, indicated that 4AF8 demonstrated the best ligand fitting (Section 4.3.5.1). As a result, both *Tco*MCA5

and *Tvi*MCA5 were modelled using *Tbb*MCA2 4AF8 and superimposed onto the template (Fig. 4.2, panel A). The generated alignment from SwissModel is accompanied by Qmean scores which are represented in blue indicating a high-quality prediction, and in orange/red indicating a low-quality prediction (Fig. 4.2, panels B and C). The alignments are enlarged in Appendix C3.

The final 3D structures of *Tco*MCA5 and *Tvi*MCA5 were validated using the protein structure analysis software, ProSA-web (Fig. 4.3) (Wiederstein and Sippl, 2007). The Z-score indicates the overall quality of the protein model and shows that the scores for the 3D structure, solved by X-ray diffraction, of *Tbb*MCA2 (-7.93) and homology modelled *Tco*MCA5 (-9.12) and *Tvi*MCA5 (-9.12) are within the range of scores (-8 to -12) found for native proteins of a similar size as determined by X-ray diffraction.

The Ramachandran plots for both *Tco*MCA5 and *Tvi*MCA5 are very similar to that of *Tbb*MCA2 4AF8, whose structure was solved at a resolution of 1.4 Å using X-ray diffraction (Fig. 4.4). Considering all the data from Fig. 4.2 to 4.4, the homology models generated for *Tco*MCA5 and *Tvi*MCA5 were suitable for docking studies with the Berg and commercial libraries.

4.3.3 Protein preparation

The N-terminal domain of *Tbb*MCA2 is a well-ordered loop, consisting of 70 residues, which encircles the main body, forming 32 hydrogen bonds and eight salt bridges, and crosses over the active site cavity (McLuskey *et al.*, 2012) (Fig. 4.5, panel A). The domain crosses over the active site cavity at Tyr31 (Fig. 4.5, panel B). A calcium-induced conformational change occurs in *Tbb*MCA2 whereby it is suggested that the N-terminal domain is removed from the entrance of the active site using Tyr31 as the gatekeeper, together with Ser156 as the latch (McLuskey *et al.*, 2012). As such, residues 1 to 35 for *Tbb*MCA2 and 1 to 25 for both *Tco*MCA5 and *Tvi*MCA5 were deleted to mimic the conformational change and allow access into the active site (Fig. 4.5, panel C). To validate the removal of the prodomain, binding site identification was carried out with and without the N-terminal domain. No binding site was identified at the active site while the N-terminal domain was present, but upon its removal, a binding site was identified in the same location of the active site.

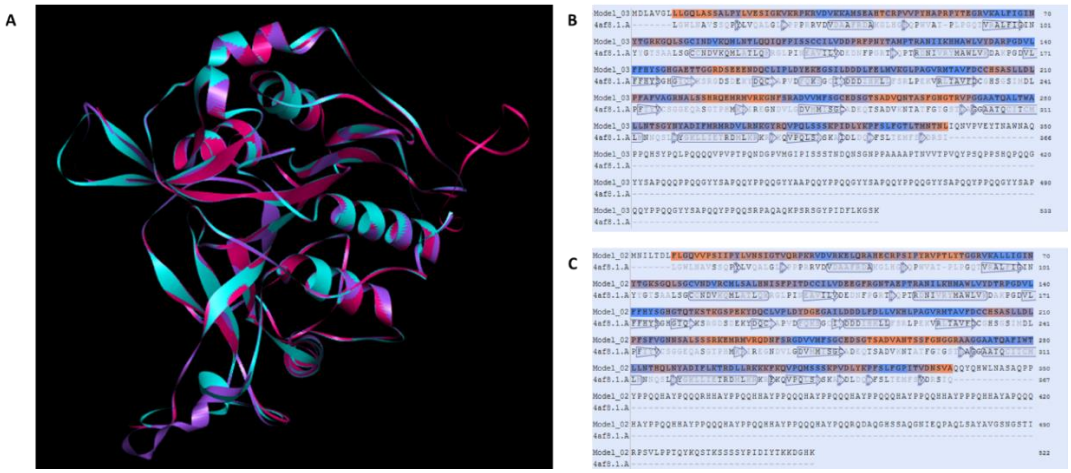


Figure 4.2: Homology modelling of *Tco*MCA5 and *Tvi*MCA5. Using *Tbb*MCA2 4AF8 (pink) as the template, the 3D structures of *Tco*MCA5 (purple) and *Tvi*MCA5 (turquoise) were **(A)** modelled using SwissModel and superimposed. The alignment of the **(B)** *Tco*MCA5 and **(C)** *Tvi*MCA5 sequences to the *Tbb*MCA2 4AF8 template is shown with the Qmean scores indicated in blue (high quality prediction) and orange/red (low quality prediction). (See Appendix C3 for enlarged versions of panels B and C).

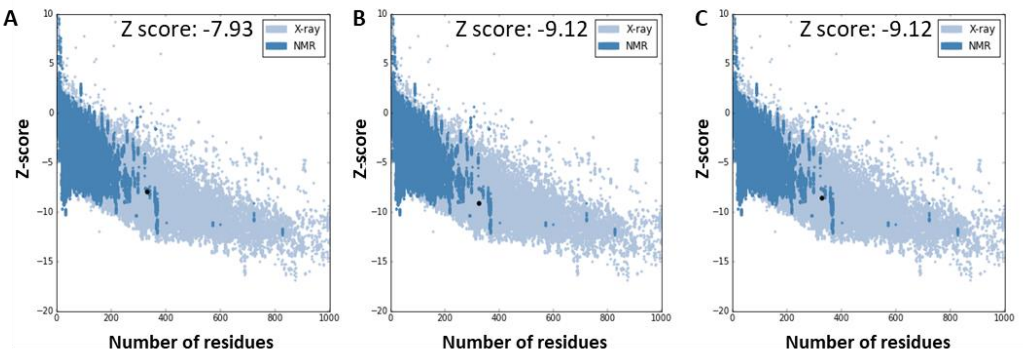


Figure 4.3: Z-score plot as an indication of model quality. The PDB files of the resulting **(A)** *Tbb*MCA2 4AF8 and homology modelled **(B)** *Tco*MCA5 and **(C)** *Tvi*MCA5 structures, were individually uploaded into the ProSA software for evaluation (Wiederstein and Sippl, 2007).

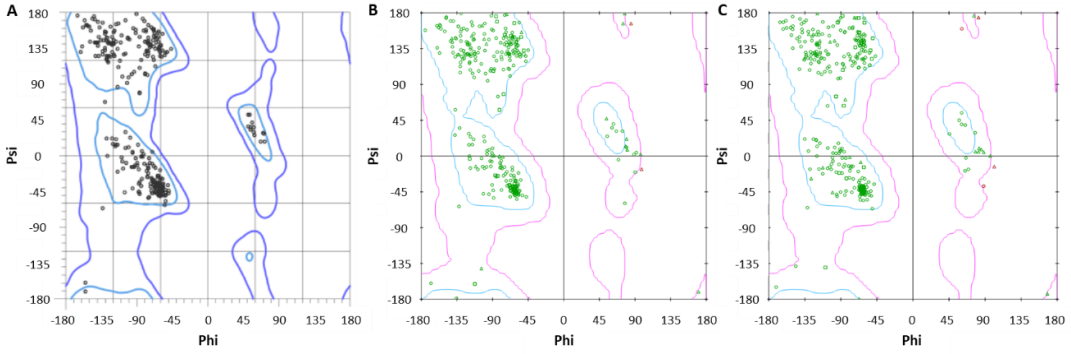


Figure 4.4: Ramachandran plots of *Tbb*MCA2 4AF8 and modelled *Tco*MCA5 and *Tvi*MCA5. The plots for **(A)** *Tbb*MCA2 4AF8 was obtained from the PDB (Berman *et al.*, 2000) (McLuskey *et al.*, 2012) and those for homology modelled **(B)** *Tco*MCA5 and **(C)** *Tvi*MCA5 were generated using the Discovery Studio software (Dassault Systèmes BIOVIA).

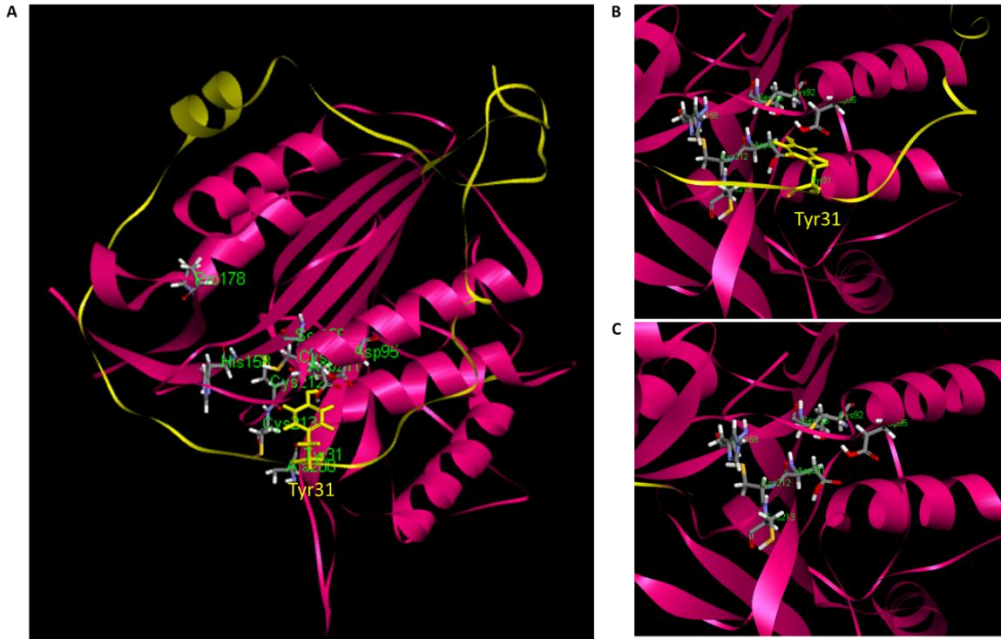


Figure 4.5: Interactions of the N-terminal domain with the main body of *Tbb*MCA2 4AF8. **(A)** The N-terminal domain (yellow) encircles the main body of *Tbb*MCA2 4AF8 (pink), **(B)** and blocks access to the active site using Tyr31. **(C)** After removal of residues 1 to 35, the active site is accessible to ligands, substrates and inhibitors.

4.3.4 Active site identification

After the deletion of the N-terminal domain residues of *Tbb*MCA2 (1 to 35) and MCA5 (1 to 25), the binding sites were identified by the cavity method using the Discovery Studio protocol (Dassault Systèmes BIOVIA).

The binding site used for docking studies needs to be located at the active site cavity which contains the catalytic dyad, S_1 binding, adjacent and secondary catalytic Cys residues (Fig. 4.6, panel A) as well as the three conserved water molecules (Fig. 4.6, panel B). The spatial positions of these residues are similar and align well between *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5.

The binding sites, which corresponded to the location of the active site in each of the four *Tbb*MCA2 models, the homology modelled *Tco*MCA5 and *Tvi*MCA5, and in each of the four water preparations, were selected and detailed in Table 4.2. The surface of the active site of *Tbb*MCA2 4AF8 in the four water preparations is depicted in Fig. 4.7.

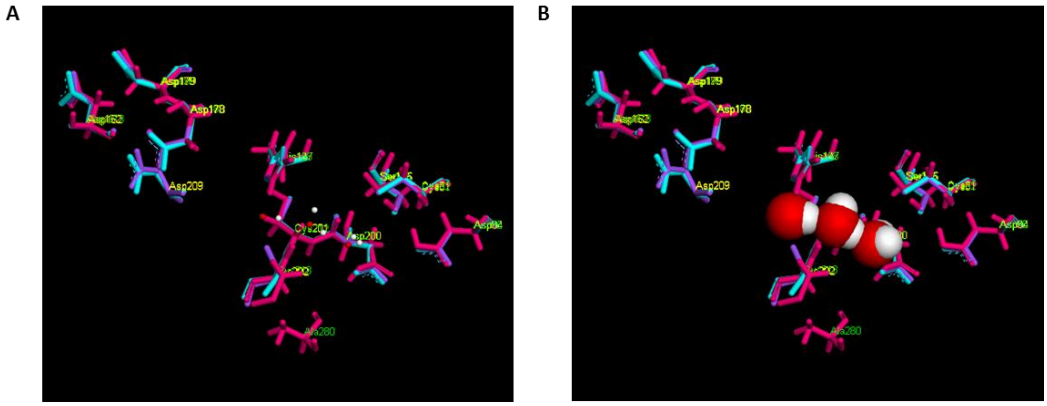


Figure 4.6: Identification of the active site residues of *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5. (A) The residues involved in calcium binding, in substrate binding and catalysis are highlighted for *Tbb*MCA2 4AF8 (pink), *Tco*MCA5 (purple) and *Tvi*MCA5 (turquoise). **(B)** Enlarged conserved water molecules lining the bottom of the active site.

The 'all water' and 'conserved water without preparation' preparations of *Tbb*MCA2 4AF8 yielded the same structure (Fig. 4.8, panel A) with identical binding site positions (Table 4.2), hence the appearance of a single structure. A noticeable difference was observed between the 'conserved water without preparation' and 'no water' (panel B) as well as with 'conserved water with preparation' (panel C) where the location of the loop C-terminal to Cys213 varies. Hence, two different coloured structures are observed in panels B and C. The coordinates of the binding sites of the 'no water' and 'conserved water prepared' preparations are different from the 'all water' and 'conserved water without preparation'. Similar structural features were obtained for both *Tco*MCA5 and *Tvi*MCA5.

Table 4.2: Binding site spheres, in which the catalytically important residues are found, to be used for docking studies.

Protein (PDB ID)	'All water'				'No water'				'Conserved water, prepared'				'Conserved water without preparation'			
	Radius (Å)	Coordinates (Å)			Radius (Å)	Coordinates (Å)			Radius (Å)	Coordinates (Å)			Radius (Å)	Coordinates (Å)		
		X	Y	Z		X	Y	Z		X	Y	Z		X	Y	Z
*Tbb*MCA2 (4AF8)	8	55.57	43.89	-5.72	6	60.1	43.67	-10.71	6	59.88	42.36	-8.63	8	58.57	43.89	-5.72
*Tbb*MCA2 (4AFP)	6	59.55	42.19	-7.04	6	60.13	43.19	-10.39	6	59.86	41.86	-8.47	6	59.55	42.19	-7.042
*Tbb*MCA2 (4AFV)	8	64.02	47.21	-12.14	10	57.78	43.41	-15.65	6	60.09	41.85	-8.40	8	64.02	47.21	-12.14
*Tbb*MCA2 (4AFR)	6	60.64	44.01	-8.48	10	57.52	43.33	-17.14	6	60.02	42.14	-8.76	6	60.64	44.01	-8.48
*Tco*MCA5	8.6	58.74	43.88	-5.62	6	60.05	43.61	-10.34	6	60.01	43.23	-8.93	6	60.01	43.54	-9.33
*Tvi*MCA5	9	57.59	43.32	5.64	6	60.04	43.60	-10.84	6	57.71	41.92	-8.50	6	59.80	42.02	-8.45

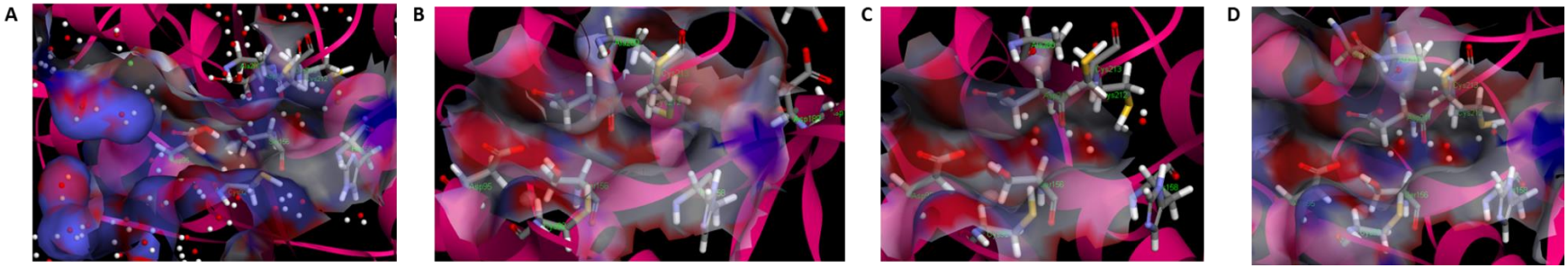


Figure 4.7: Active site surface of *Tbb*MCA2 4AF8 after deletion of residues 1 to 35 of the N-terminal domain. The surface of the active site cavity was determined in the various water preparations, **(A)** 'all water', **(B)** 'no water', **(C)** 'conserved water, prepared' and **(D)** 'conserved water without preparation'. The atomic structures of the catalytic dyad (His158, Cys213), adjacent Cys (Cys212), secondary catalytic Cys (Cys92) and S_1 binding residues (Asp95, Ser156 and Asp211) are represented in a stick configuration.

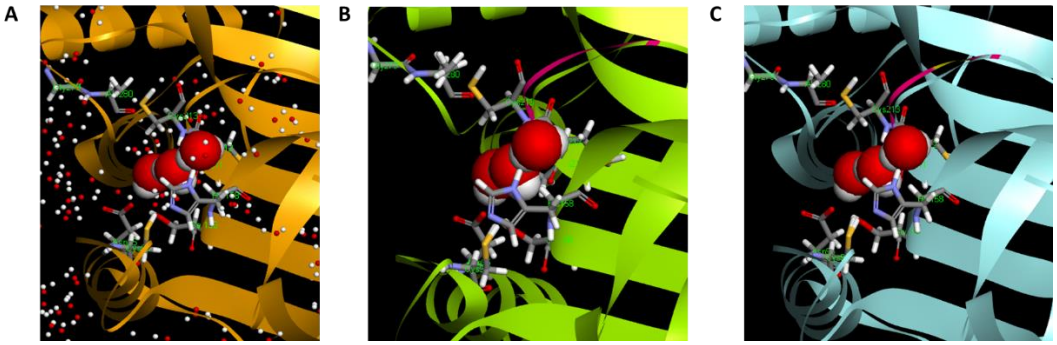


Figure 4.8: Structural differences between the different water preparations of *Tbb*MCA2 4AF8. Overlay of **(A)** 'all water' (light orange) and 'conserved water without preparation' (pink), **(B)** 'no water' (light green) and 'conserved water without preparation' (pink), and **(C)** 'conserved water prepared' (blue-grey) and 'conserved water without preparation' (pink).

4.3.5 CDocker analysis

The substrates and inhibitors of the commercial library were shown in literature to be cleaved by MCAs and inhibit the MCAs, respectively. Together with the ligands in the Berg library, which were designed based on the P_1 specificity of *Tbb*MCA2, the substrates and inhibitors of the commercial library were docked into the active site of *Tbb*MCA2 4AF8, 4AFP, 4AFV, 4AFR, *Tco*MCA5 and *Tvi*MCA5, in each of the four water preparations.

The CDocker analysis predicts the interaction energy associated with each ligand conformation and which ligands are docked in the specified binding site (Kapetanovic, 2008). Ligands with a high -CDocker interaction energy (kcal/mol) possess a higher binding affinity for the protein 'receptor' than those with lower energies. Multiple poses for each ligand are generated, with only the pose with the highest -CDocker interaction energy being reported. The IC_{50} values reported in the analyses are of the ligands tested against recombinant *Tbb*MCA2 as published by Berg *et al*. (2012). The correlation of -CDocker interaction energies of the ligands, substrates and inhibitors docked into the active site of *Tbb*MCA2*, Tco*MCA5 and *Tvi*MCA5 together with the IC_{50} of the Berg ligands are given in Appendix C4.

4.3.5.1 *Tbb*MCA2

The -CDocker interaction energies for each ligand/inhibitor/substrate, which were successfully docked into the active site of *Tbb*MCA2, from the commercial and Berg libraries are reported in Table 4.3 and 4.4, respectively. Since the substrates and inhibitors of the commercial library have been shown experimentally in literature to be

cleaved by MCAs, or inhibit the MCAs, these results were used to select the best water preparation for the analysis of the docking results using the Berg library.

The 3D structure, solved by X-ray diffraction at a resolution of 1.4 Å, of *Tbb*MCA2 4AF8 had the best resolution compared to the other three *Tbb*MCA2 structures, 1.6 and 2.1 Å. Overall, more of the ligands from the commercial and Berg libraries were docked into the active site of 4AF8, with higher -CDocker interaction energies, than compared to those docked in the other *Tbb*MCA2 structures.

Out of the four water preparations of *Tbb*MCA2 4AF8, only the 'all water' and 'conserved water without preparation' preparations, docked all the ligands with higher -CDocker interaction energies than the other two water preparations. This may be a direct effect of the difference in structure as shown in Fig. 4.8. As such, only the results for the 'all water' preparation will be analysed.

Antipain had the highest -CDocker interaction energy, 73 kcal/mol, of the four inhibitors, followed by chymostatin, 65 kcal/mol, and leupeptin, 53 kcal/mol. Low -CDocker interaction energies of 38 and 28 kcal/mol were obtained for TLCK and AEBSF, respectively.

The fluorogenic peptide substrates which had the highest affinity for *Tbb*MCA2 4AF8 were Boc-Val-Pro-Arg-AMC, 3-Methoxysuccinyl (MeOSuc)-Ala-Phe-Lys-AMC and Z-Gly-Gly-Arg-AMC with -CDocker interaction energies of 78, 71 and 69 kcal/mol, respectively. The remaining substrates had -CDocker interaction energies between 68 and 61 kcal/mol. The substrates used as active site inhibitors, Phe-Arg-FMK (Klemenčič *et al.*, 2015) and Phe-Pro-Arg-chloromethylketone (CMK) (Watanabe and Lam, 2011) had the lowest -CDocker interaction energies of 59 and 54 kcal/mol, clearly different from the fluorogenic peptide substrates.

The Berg ligands with low IC_{50} values generally had higher -CDocker interaction energies than those with IC_{50} values above 100 µM. This was more evident for the 'conserved water without preparation' when compared to the 'all water' preparation model.

Table 4.3: The -CDocker interaction energies for the inhibitors and substrates of the commercial library which were successfully docked into the active site, of the four water preparations, of *Tbb*MCA2 4AF8.

'All water'		'No water'		'Conserved water, prepared'		'Conserved water without preparation'	
Ligand	Interact. energy[a]	Ligand	Interact. energy[a]	Ligand	Interact. energy[a]	Ligand	Interact. energy[a]
Boc-Val-Pro-Arg-AMC	77.9998	TLCK	46.1507	TLCK	39.3022	Antipain	73.4214
Antipain	73.8336	AEBSF	37.7071	AEBSF	30.6874	MeOSuc-Ala-Phe-Lys-AMC	65.2483
MeOSuc-Ala-Phe-Lys-AMC	71.4556	Phe-Pro-Arg-CMK	31.8880			Z-Gly-Gly-Arg-AMC	64.1411
Z-Gly-Gly-Arg-AMC	69.4528					Val-Leu-Arg-AMC	63.8513
Val-Leu-Lys-AMC	68.0078					Chymostatin	63.1609
Val-Leu-Arg-AMC	66.7398					Boc-Val-Pro-Arg-AMC	62.8661
Chymostatin	65.4091					Val-Leu-Lys-AMC	61.9325
Phe-Arg-FMK	59.5988					Phe-Arg-FMK	54.6182
Phe-Pro-Arg-CMK	54.1384					Phe-Pro-Arg-CMK	51.2296
Leupeptin	53.6793					Leupeptin	50.5274
TLCK	38.8368					TLCK	39.1133
AEBSF	28.0329					AEBSF	29.0082

[a]The interaction energy is reported as kcal/mol.
Substrates and inhibitors were individually colour-coded for ease of interpretation of data.

Table 4.4: The -CDocker interaction energies for the ligands of the Berg library which were successfully docked into the active site, of the four water preparations, of *Tbb*MCA2 4AF8.

'All water'			'No water'			'Conserved water, prepared'			'Conserved water without preparation'		
Ligand name	IC_{50} (μM)[b]	Interact. energy[a]	Ligand	IC_{50} (μM)[b]	Interact. energy[a]	Ligand name	IC_{50} (μM)[b]	Interact. energy[a]	Ligand name	IC_{50} (μM)[b]	Interact. energy[a]
777	1.4	57.3327	297	1.6	49.4469	773	2.2	44.5092	777	1.4	55.1019
778	1.1	56.6135	772	62.0	47.5740	297	1.6	41.9132	778	1.1	54.6875
775	3.9	49.0600	773	2.2	46.6269	779	1.9	40.8882	774	38.4	45.0717
772	62.0	48.7148	781	>100	46.1817	772	62.0	39.6582	775	3.9	44.8550
776	4.0	48.6662	774	38.4	43.7821	771	0.6	35.9593	776	4.0	44.6914
771	0.6	48.5409	776	4.0	42.7313	765	>100	35.1611	772	62.0	44.6729
770	>100	48.5346	779	2	42.2552	781	>100	34.9746	771	0.6	44.4879
774	38.4	48.2427	767	>100	41.7179	782	>100	33.2580	297	1.6	43.8606
769	>100	46.8011	771	0.6	41.6451	780	>100	32.6337	770	>100	43.7727
297	1.6	46.5066	782	>100	41.3783	767	>100	31.2158	773	2.2	42.3363
773	2.2	42.8317	768	>100	40.9039	766	>100	30.4471	781	>100	40.2067
767	100	39.9584	765	>100	40.7637				779	1.9	40.1646
779	1.9	38.9667	766	>100	38.8653				769	>100	38.3082
781	>100	38.3100	775	3.9	35.0394				768	>100	33.9735
768	>100	37.5578	780	>100	34.7704				765	>100	33.4806
766	>100	37.3704							767	>100	33.0304
780	>100	36.3873							780	>100	32.5654
765	>100	34.9376							766	>100	32.5549
782	>100	32.5698							782	>100	28.7093

[a]The interaction energy is reported as kcal/mol.
[b]The IC_{50} values were obtained from Berg *et al*. (2012).
Ligands were individually colour-coded for ease of interpretation of data.

The commercial inhibitor, antipain, had the highest -CDocker interaction energy and has been shown to inhibit recombinant *Tbb*MCA2 (Moss *et al.*, 2007) and recombinant *Tco*MCA5 in the results reported in the Chapter 3. The spatial position of the inhibitor, substrate and ligand with the highest -CDocker interaction energies in the active site of *Tbb*MCA2 4AF8, with the 'all water' preparation, is shown in Fig. 4.9. The 2D views and residue interactions are reported in Appendix C5. Despite the Z-Gly-Gly-Arg-AMC substrate having the third highest -CDocker interaction energy, it is the most commonly used substrate for the characterisation of MCAs in the published literature, and was used in the work described in Chapter 3 for the characterisation of recombinant *Tco*MCA5. The docked position of Z-Gly-Gly-Arg-AMC and the Berg ligand 777, which has a low IC_{50} value (1.4 μM) is shown in Fig. 4.9.

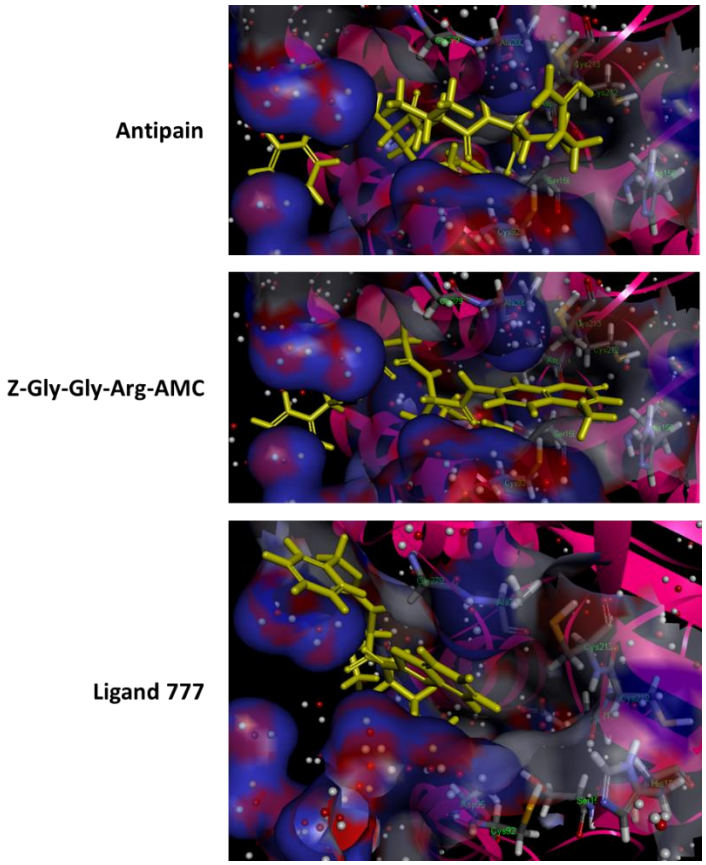


Figure 4.9: The commercial inhibitor, substrate and Berg ligand which had the highest binding affinity, at the active site, for *Tbb*MCA2 4AF8, prepared with 'all water.' The 3D interactions of the **(A)** commercial inhibitor, antipain, **(B)** commercial substrate, Z-Gly-Aly-Arg-AMC and **(C)** Berg ligand, 777, docked in the active site of *Tbb*MCA2 4AF8.

4.3.5.2 *Tco*MCA5 and *Tvi*MCA5

The docking of the commercial and Berg libraries into the active sites of *Tco*MCA5 and *Tvi*MCA5, which had been modelled on *Tbb*MCA2 4AF8, using all four water

preparations, showed a similar result to what was seen for *Tbb*MCA2 4AF8. Thus, only the results from the 'all water' preparation of *Tco*MCA5 and *Tvi*MCA5 is reported here.

The -CDocker interaction energies for both the docked commercial and Berg ligands were slightly lower than those of *Tbb*MCA2 4AF8, but overall, were very similar. (Table 4.5). Antipain was the best docked inhibitor, followed by chymostatin, with a 15 kcal/mol difference between antipain and leupeptin. Lower -CDocker interaction energy, between 40 to 23 kcal/mol was observed between *Tco*MCA5 and *Tvi*MCA5 with TLCK and AEBSF.

Similar to *Tbb*MCA2 4AF8, the best docked substrate onto *Tco*MCA5 was MeOSuc-Ala-Phe-Lys-AMC and Boc-Val-Pro-Arg-AMC for *Tvi*MCA5. The Z-Gly-Gly-Arg-AMC substrate had the third best -CDocker interaction energy for *Tco*MCA5 and the fifth best for *Tvi*MCA5, with a 6 and 9 kcal/mol difference from that of the best docked substrate. The substrates reported in literature as active site inhibitors, Phe-Arg-FMK (Klemenčič *et al.*, 2015) and Phe-Pro-Arg-CMK (Watanabe and Lam, 2011) docked onto *Tco*MCA5 and *Tvi*MCA5 and had almost the exact same -CDocker interaction energies to those docked onto *Tbb*MCA2 4AF8 (Table 4.3).

Based on the results for *Tbb*MCA2 4AF8, the position of antipain, Z-Gly-Gly-Arg-AMC and Berg ligand 777, which docked with either the highest or second highest -CDocker interaction energies in the active site of *Tco*MCA5 and *Tvi*MCA5, with the 'all water' preparation, is shown in Fig. 4.10. The 2D views and residue interactions are reported in Appendix C6.

Overall, similar -CDocker interaction energies were obtained for the ligands of the Berg library when docked into the active sites of *Tbb*MCA2 4AF8, *Tco*MCA5 and *Tvi*MCA5 with the 'all water' preparation. The ligands which had the highest -CDocker interaction energies were 777 and 778. The ligands with high IC_{50} values (>100 µM), 770 and 769, had higher -CDocker interaction energies than 779 which had a low IC_{50} value (1.9 µM) in the above three protein 'receptor' models.

Table 4.5: The -CDocker interaction energies for the inhibitors and substrates from the commercial library, and ligands from the Berg library, which were successfully docked into the active site, of the 'all water' preparation, of *Tco*MCA5 and *Tvi*MCA5.

Commercial library				**Berg library**					
*Tco*MCA5		*Tvi*MCA5		*Tco*MCA5			*Tvi*MCA5		
Ligand name	Interact. energy[a]	Ligand name	Interact. energy[a]	Ligand name	IC_{50} (μM)[b]	Interact. energy[a]	Ligand name	IC_{50} (μM)[b]	Interact. energy[a]
Antipain	69.3579	Boc-Val-Arg-Pro-Arg-AMC	69.0750	777	1.4	56.7501	777	1.4	58.0676
MeOSuc-Ala-Phe-Lys-AMC	69.1403	Antipain	69.0354	778	1.1	56.5844	778	1.1	54.8880
Chymostatin	68.7365	MeOSuc-Ala-Phe-Lys-AMC	68.1931	772	62.0	50.7902	775	3.9	50.0416
Val-Leu-Arg-AMC	65.779	Val-Leu-Arg-AMC	66.5845	770	>100	50.7726	776	4.0	49.9786
Z-Gly-Gly-Arg-AMC	63.9770	Chymostatin	66.3283	774	38.4	49.4546	770	>100	48.1789
Boc-Val-Arg-Pro-Arg-AMC	61.7294	Val-Leu-Lys-AMC	62.1262	775	3.9	49.1105	772	62.0	47.8103
Phe-Arg-FMK	59.6302	Z-Gly-Gly-Arg-AMC	60.4820	771	0.6	48.7838	774	38.4	47.6948
Val-Leu-Lys-AMC	59.2276	Phe-Arg-FMK	57.5471	776	4.0	47.5925	771	0.6	46.8969
Leupeptin	54.5873	Phe-Pro-Arg-CMK	54.9999	769	>100	46.1706	773	2.2	45.5624
Phe-Pro-Arg-CMK	54.2661	Leupeptin	54.2910	297	1.6	46.0371	769	>100	43.6866
TLCK	40.0068	TLCK	37.2190	773	2.2	45.9522	297	1.6	43.2771
AEBSF	25.8711	AEBSF	23.9091	768	>100	43.6187	766	>100	39.1067
				781	>100	41.7428	782	>100	38.2214
				767	>100	40.1867	779	1.9	38.0768
				766	>100	39.6047	780	>100	37.5859
				779	1.9	39.3659	765	>100	37.5521
				765	>100	38.0949	768	>100	37.1429
				780	>100	35.7624	767	>100	36.6743
				782	>100	34.9732	781	>100	36.2713

[a]The interaction energy is reported as kcal/mol.
[b]The IC_{50} values were obtained from Berg *et al*. (2012).
Substrates, inhibitors and ligands were individually colour-coded for ease of interpretation of data.

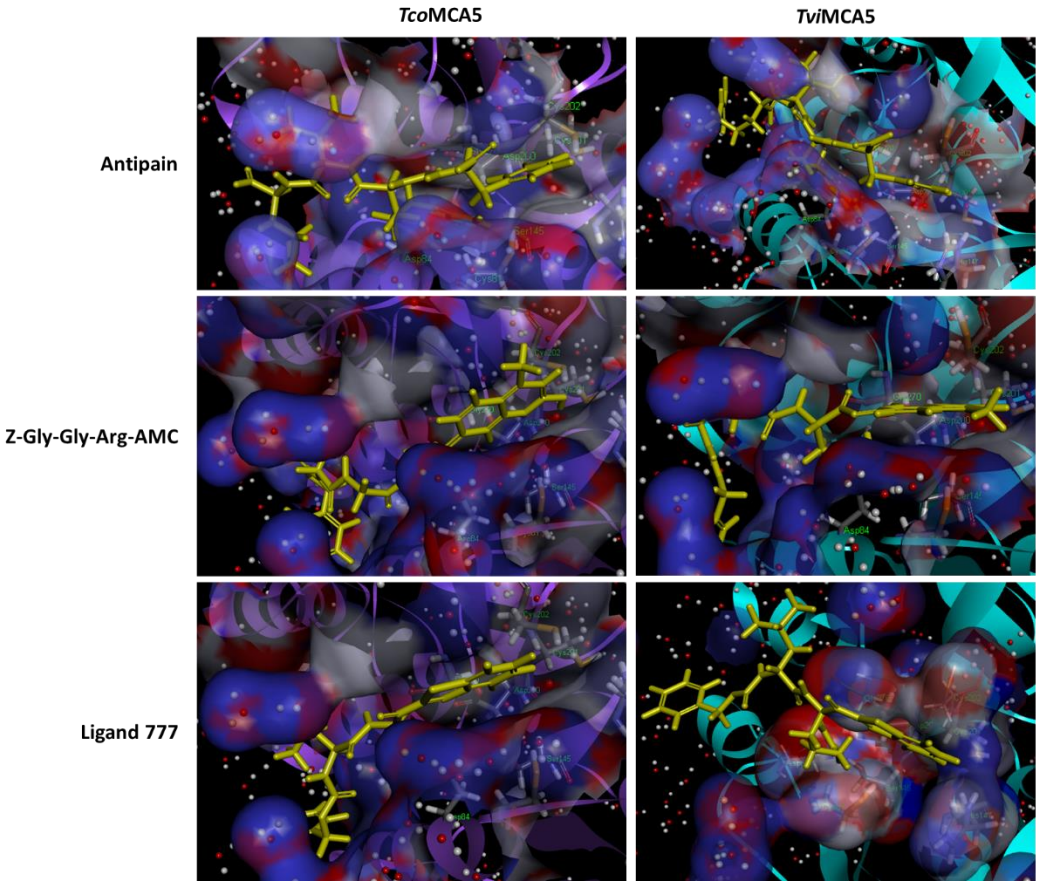


Figure 4.10: The inhibitor, substrate and Berg ligand which had the highest binding affinity at the active site for *Tco*MCA5 and *Tvi*MCA5, prepared with 'all water'. The 3D and 2D interactions of the commercial inhibitor, antipain, commercial substrate, Z-Gly-Gly-Arg-AMC and Berg ligand, 777 docked in the active site of *Tco*MCA5 and *Tvi*MCA5.

4.3.6 Ligand interactions

Overall, more interactions between the ligands and the MCA 'receptor' were formed between the inhibitors and substrates of the commercial library than what was formed with the ligands from the Berg library (Fig. 4.11). Prominent residues, with which the library ligands formed hydrogen bonds (panel A) with *Tbb*MCA2 4AF8 (equivalent residues in *Tco*MCA5 and *Tvi*MCA5 in brackets) were Ser90 (Ser79), Asp95 (Asp84), Gly279 (Val268), Ala280 (Pro269), Gly281 (Gly270) and Lys314 (Lys304).

Charge interactions (panel B) of the residues of *Tbb*MCA2 4AF8 (equivalent residues in *Tco*MCA5 and *Tvi*MCA5 in brackets) were most prominently formed with Asp95 (Asp84) with the commercial ligands. Fewer interactions were formed with Asp84 of *Tco*MCA5 and *Tvi*MCA5, but an increase in Lys303 and Tyr305 interactions was evident (Lys314 and Phe316 in *Tbb*MCA2).

A noticeable difference in the number of hydrophobic interactions between the MCA 'receptors' and the ligands were observed, with more interactions being formed with those of the commercial library than with those of the Berg library (panel C). Residues of the *Tco*MCA5 and *Tvi*MCA5 (equivalent residues in *Tbb*MCA2 in brackets) involved were Leu89 (Leu78), secondary catalytic Cys81 (Cys92), catalytic Cys202 (Cys213), Val268, Pro269 (*Tco*MCA5 only) and Tyr305 (Phe316).

Pi-alkyl and Pi-sulfur interactions were more prominent in *Tco*MCA5 and *Tvi*MCA5 than in *Tbb*MCA2 4AF8 (equivalent residues in *Tbb*MCA2 in brackets), with interactions occurring at the secondary catalytic Cys81 (Cys92), and the catalytic Cys202 (Cys213) (panel D).

The interactions formed, type and frequency, between ligand and 'receptor' highlight key features to take into account when considering the design of novel inhibitors or adjustment of the current ligands to improve binding affinity when combined with a common feature pharmacophore.

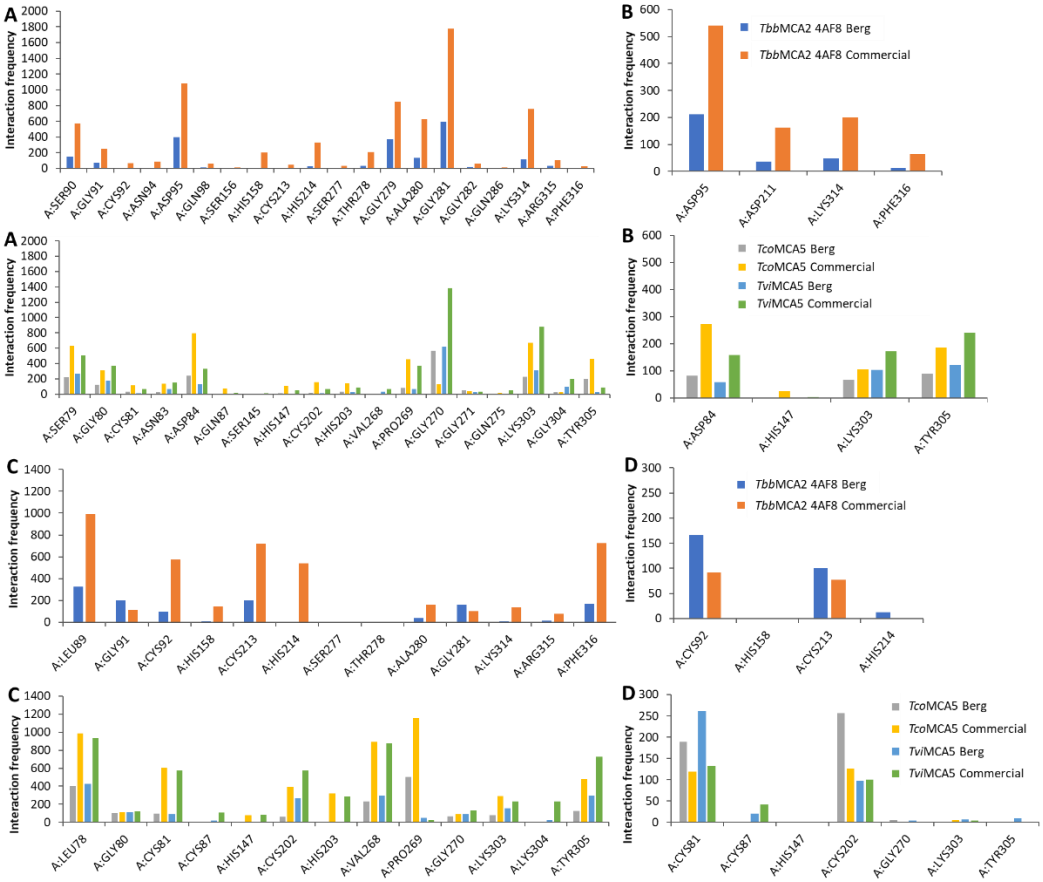


Figure 4.11: Interactions of inhibitors, substrates and ligands of the commercial and Berg libraries with *Tbb*MCA2 4AF8, *Tco*MCA5 and *Tvi*MCA5. The interaction frequency between residues and ligands in terms of **(A)** hydrogen bond formation, **(B)** charged interactions, **(C)** hydrophobic interactions and **(D)** Pi-alkyl and Pi-sulfur.

4.3.7 Common feature pharmacophore

To facilitate the selection of inhibitors/substrates/ligands to use in the generation of a common feature pharmacophore, the -CDocker interaction energies as well as the adopted conformations of the ligands in the active site were considered. Since the inhibitors and substrates from the commercial library have been successfully used in literature for the characterisation of various MCAs, the best aligned substrates, with as high -CDocker interaction energies as possible, were used. The commercial inhibitors were not used for the generation of the pharmacophore as they adopted very different conformations in the active site, and as such, did not align reasonably well, when compared to the commercial substrates and Berg ligands (Fig. 4.12).

The commercial substrates, MeOSuc-Ala-Phe-Lys-AMC, Boc-Val-Pro-Arg-AMC, Z-Gly-Gly-Arg-AMC and Val-Leu-Arg-AMC, were docked in similar conformations compared to the remaining substrates which did not adopt similar conformations in the active site. These four substrates were used to generate a common feature pharmacophore for *Tbb*MCA2 4AF8, *Tco*MCA5 and *Tvi*MCA5 (Fig. 4.13).

The pharmacophore for *Tbb*MCA2 4AF8 is characterised by one positive ionisation and three hydrophobic features, together with one hydrogen bond donor and two hydrogen bond acceptors. For *Tco*MCA5, the pharmacophore consists of one ionisation and three hydrophobic features, together with three hydrogen bond acceptors. Four hydrogen bond acceptors as well as one positive ionisation and three hydrophobic features, constitute the pharmacophore of *Tvi*MCA5.

4.4 Discussion

An important component of modern medicinal chemistry is the use of the 3D structure of protein targets in structure-based drug design (SBDD) methods (Salum *et al.*, 2008). Lounnas *et al*. (2013) proposed that the combination of homology modelling, molecular docking and virtual screening could be used to accelerate drug discovery. Njogu *et al*. (2016) contended that predictions using molecular docking can be improved using the experimental data from phenotypic whole-cell assays toward drug discovery for malaria, tuberculosis, trypanosomosis, and leishmaniasis.

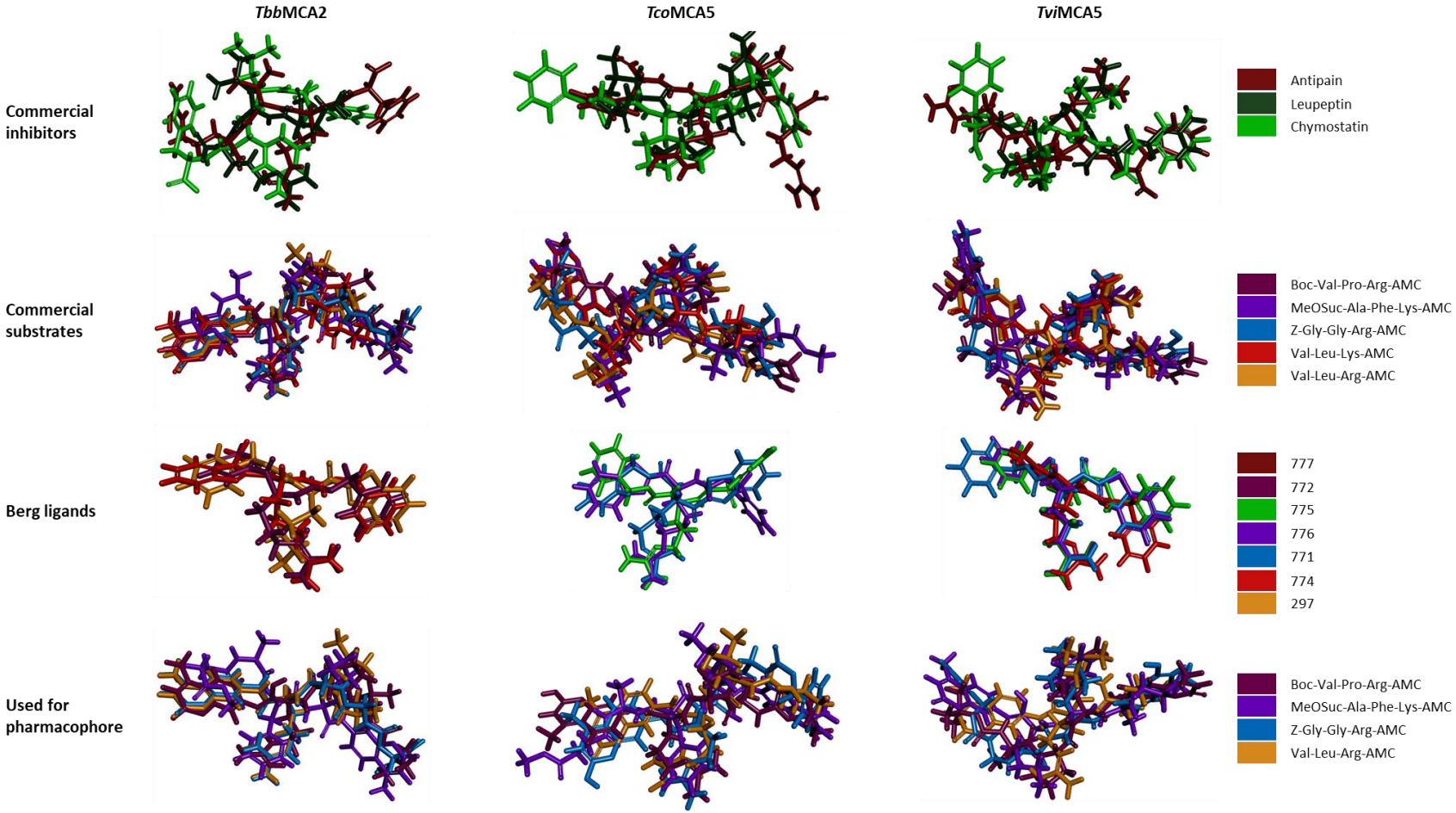


Figure 4.12: Overlay of docked commercial inhibitors and substrates along with Berg ligands into the active site of *Tbb*MCA2 4AF8, *Tco*MCA5 and *Tvi*MCA5, prepared with 'all water'.

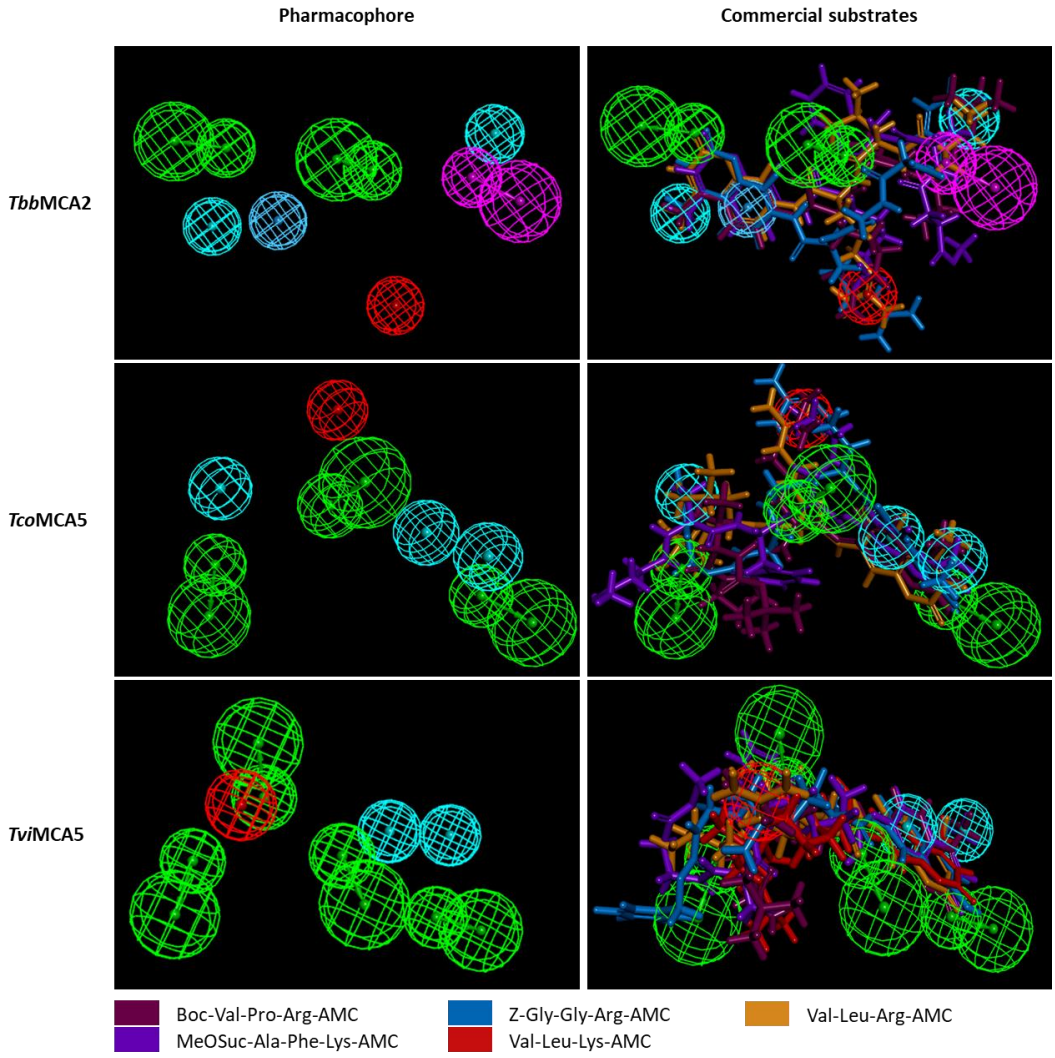


Figure 4.13: Generation of a common feature pharmacophore for *Tbb*MCA2 4AF8, *Tco*MCA5 and *Tvi*MCA5. The features of the generated pharmacophore using the docked commercial substrates are coloured as follows: hydrogen bond acceptor (green), hydrogen bond donor (pink), positive ionisation (red) and hydrophobic (turquoise).

Molecular docking can provide insight into the conformation adopted by small-molecule ligands within the binding site of a protein 'receptor' (Meng *et al.*, 2011). Using the structural alignment of a set of docked small-molecule ligands, a pharmacophore can be created (Lounnas *et al.*, 2013). The application of molecular docking studies have enabled investigations into ligand−'receptor' interactions as well as the elucidation of the mechanism of action (Njogu *et al.*, 2016). Together with this information, SBDD can direct the choice or design of ligands possessing the required features to control the activity of the target 'receptor' (Mandal *et al.*, 2009; Blaney, 2012).

More recently, phenotypic screening against *T. b. brucei*, *T. cruzi* and *L. donovani* allowed for the assembly of three anti-kinetoplastid chemical boxes targeting *T. b. brucei*, *T. cruzi* and *Leishmania donovani*. (Peña *et al.*, 2015). Inhibitors of

cruzipain, a cathepsin-L like peptidase of *T. cruzi*, were identified within the *T. b. brucei* HAT and *T. cruzi* Chagas boxes (Salas-Sarduy *et al.*, 2017). Common scaffolds of the inhibitors were identified along with predicted conformations within the active site and the interactions of the inhibitor with the cruzipain 'receptor' (Salas-Sarduy *et al.*, 2017).

Trypanosomatid cysteine peptidases, such as cruzipain and rhodesian, the cathepsin-L-like peptidase from *T. b. rhodesiense* (Mendez-Lucio *et al.*, 2012), have been validated as drug targets. Through the utilisation of structure-based screening, a number of promising compounds have been identified (Ferreira and Andricopulo, 2017). In addition, glyceraldehyde-3-phosphate dehydrogenase (Bressi *et al.*, 2001), trypanothione reductase (Maccari *et al.*, 2011) and tubulin (Goodarzi *et al.*, 2010) have been identified as drug targets using both structure- and ligand-based drug design methodologies.

The MCA cysteine peptidases are virulence factors and have been implicated in cell death (Lee *et al.*, 2007; Zalila *et al.*, 2011) and cell cycle regulation (Ambit *et al.*, 2008; Laverrière *et al.*, 2012) in the kinetoplastid parasites. The 3D structure, solved by X-ray diffraction, of *Tbb*MCA2 (McLuskey *et al.*, 2012) was elucidated two months after the creation of the Berg library of ligands which was designed based on peptidomimetics (2012). Using molecular docking studies, the effectiveness of the Berg ligands against both *Tbb*MCA2 and the MCA5 of *T. congolense* and *T. vivax*, was investigated in the present study. In addition to the Berg library, a commercial library was assembled based on the peptide substrates and inhibitors used to characterise MCAs as reported in literature.

Homology models are reasonable docking templates, provided that that they are built from a template with which it shares a high sequence similarity (Schmidt *et al.*, 2014). Homology modelling of α-glucosidase, and the subsequent docking of stilbene urea derivatives into the active site resulted in the identification of interactions required for high binding affinity (Lee *et al.*, 2014). Due to the limited numbers of solved MCA structures and the 50% sequence identity between *Tbb*MCA2 and the MCA5s of *T. congolense* and *T. vivax*, the homology models for *Tco*MCA5 and *Tvi*MCA5 were generated using *Tbb*MCA2 4AF8 as the template.

The MCA5 homology models lacked the Pro-, Gln-, Tyr-rich C-terminal domain, but due to the fact that only the catalytic domain is active, the model would suffice for docking studies. It was determined that autoprocessing of recombinant *Tbb*MCA2 at Lys55 and Lys268 to release the catalytic domain, was not an absolute requirement for activity as mutants of these residues were still catalytically active (Gilio *et al.*, 2017).

Another factor to be taken into consideration is the conformational change which results in the destabilisation of the first 15 residues of the N-terminal domain upon calcium activation as well as in the presence of a substrate (McLuskey *et al.*, 2012). It has also been suggested that calcium has a structural effect on *Tbb*MCA2 in the vicinity of the S_2 binding pocket (Machado *et al.*, 2013). It was shown in Chapter 3 that neither the recombinant *Tco*MCA5, with both the catalytic dyad residues mutated, nor the native *Tco*MCA5 required calcium for peptidolytic activity. Despite the differences between the structures of *Tbb*MCA2 and the MCA5s, the high Qmean scores of the homology models, indicating a high-quality prediction, together with Z-scores similar to those found for native proteins of a similar size, the homology modelled *Tco*MCA5 and *Tvi*MCA5 structures were suitable for docking studies.

Deeply buried structural water molecules, or those within the active site, are highly ordered and detected in 3D structures which have been solved by X-ray diffraction (Ross *et al.*, 2012). Structural data from X-ray crystallography, at a resolution better than 2.0 Å, are able to reliably discern water molecules which are present (Sousa *et al.*, 2013). The four 3D structures of *Tbb*MCA2 were determined at a resolution of between 1.4 and 2.1 Å, and as such, the water molecules identified within the structure are reliable.

Ligand binding to the protein 'receptor' requires rearrangement of the water molecules which surround both the ligand and the 'receptor' (Lounnas *et al.*, 2013). In many cases, ligand binding is mediated by structural water molecules (Ross *et al.*, 2012). These water molecules adopt a highly ordered structure within the protein and mediate the binding of ligands to the 'receptor' through the formation of hydrogen bonds (Bissantz *et al.*, 2010; Ross *et al.*, 2012). In the case of *Lmj*OPB, two well-ordered water molecules were identified, one of which was found to coordinate with Glu621 which is responsible for the P_1 substrate specificity, and the other with Asp662 which coordinates the catalytic His697 (McLuskey *et al.*, 2010).

Strongly bound water molecules are often conserved across multiple crystallographic structures (Ferreira *et al.*, 2015), which is observed for the three water molecules lining the bottom of the active site in all four *Tbb*MCA2 structures (McLuskey *et al.*, 2012). In the active site pocket, these conserved water molecules interact with the secondary catalytic Cys92, catalytic His158, S_1 binding residues Asp211, Gly157 and Gly159. The catalytic His237 is followed by Gly238 (caspase-1 numbering) and, together with the catalytic Cys, the backbone nitrogens are involved in the formation of the oxyanion hole (Walker *et al.*, 1994; Fuentes-Prior and Salvesen, 2004). This residue (Gly159 in

*Tbb*MCA2) is conserved in each of the MCAs with the expection of *Tcr*MCA3. In addition, there is a conserved Gly (Gly158 in *Tbb*MCA2) preceding the catalytic His and may be involved in the formation of the oxyanion site. The Cys92, His158 and Asp211 were demonstrated to interact with the ligands in both the Berg and commercial libraries. This supports the notion that these conserved water molecules mediate ligand binding.

It has been demonstrated that docking performance is improved by the inclusion of water molecules in both predicted structures and those solved by X-ray diffraction (Yang and Chen, 2004; de Graaf *et al.*, 2005). A such, molecular docking studies were performed with four different water preparations. The 'all water' and 'conserved water without preparation' preparations resulted in more docked compounds with greater CDocker interaction energies than those docked in the 'no water' and conserved water with preparation'. The CDocker values obtained from the molecular docking of the Berg ligands into the active site of *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5, despite some outliers, followed the trend of the IC_{50} values obtained in the phenotypic screening (Berg *et al.*, 2012).

Reversible cruzipain inhibitors which possessed a nitrile warhead were characterised by IC_{50} values of 1 nM against the recombinant enzyme and between 5 and 10 μM in epimastigote and intracellular amastigote *in vitro* assays (Ndao *et al.*, 2014). It was expected that the Berg ligands which possessed a nitrile warhead would have a high affinity for *Tbb*MCA2 indicated by high CDocker interaction energies. This however, was not the case as shown by the IC_{50} values as well as by the low CDocker interaction energies. This may be due to the MCAs' resistance to the E-64 inhibitor (Appendix B8). This deviance from the cysteine peptidase characteristic follows the result where the inhibitor of serine peptidases (ISP), and not the inhibitor of cysteine peptidases (ICP), was shown to bind and inhibit the peptidolytic activity of the single MCA from *L. amazonensis* (Peña *et al.*, 2017). In addition, *Ldn*MCA activity was reported to be 'trypsin-like' (Lee *et al.*, 2007), which is a trait of serine peptidases rather than that of cysteine peptidases.

The substrates and inhibitors of the commercial library were successfully docked into the active site of *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5, but the serine inhibitors had higher interaction energies than those of the cysteine inhibitors. The docking results suggest that whilst antipain has been shown as the most effective inhibitor of recombinant *Tbb*MCA2 (Moss *et al.*, 2007), chymostatin was predicted to be more effective than leupeptin and TLCK which is contradictory to what is reported in literature

(Moss *et al.*, 2007). In the case of the MCA5s, the only effective inhibitors, were antipain and leupeptin, which only reduced the peptidolytic activity of recombinant *Tco*MCA5^{H147AC202G} activity by 20% (Chapter 3). Docking studies suggest that inhibition should have been observed with chymostatin as well as TLCK. Despite the differences, the ligands of the commercial library bound in a similar spatial position within the active site of *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5 which correlates to the spatially conserved active site residues (Appendix B5).

The poses are ranked according to their CDocker interaction energies where ligands with a lower associated energy possessed a higher affinity for the protein receptor. However, the calculation of docking scores and prediction of binding poses is not 100% accurate (Lounnas *et al.*, 2013). This study did not take into account the flexibility of the protein receptor upon ligand binding which affects the accuracy of the docking scores. This can be corrected using molecular dynamics, whereby protein receptor flexibility is taken into consideration during docking, which is beyond the scope of this research.

Docking of libraries containing diverse compounds are characterised by reduced reliability of scoring functions and inaccurate docking poses, thus making analysis difficult (Kitchen *et al.*, 2004). Docking scores obtained using large sets of ligands, with the same scaffold, on a series of protein kinases indicated that the correct binding mode within the top set of docked poses was obtained compared to docking 100 poses generated from a single ligand (Chema *et al.*, 2004). The Berg library and the commercial library contained compounds which were structurally similar. As such, the CDocker scores and poses should be reasonably accurate.

Without prior knowledge or a strong hypothesis of bound conformations of ligands and spatial alignment, docking scores are unreliable (Spitzer and Jain, 2012). If the preferred binding mode of a series of chemically related compounds can be inferred, docking can provide correct ligand conformation and accurate structural alignment that is required to build predictive quantitative structure–activity relationship models (Kulkarni *et al.*, 2007). Given the P_1 specificity for Arg and the favourable alignment of the docked commercial substrates and inhibitors, which have been used in previously reported MCA characterisation, the docked poses of both the Berg and commercial library can be considered to be reasonably accurate.

Due to the conservation of the caspase-hemoglobinase fold between the MCAs, the information from one structure can be transferred to others as is the case with protein kinases (Lounnas *et al.*, 2013). Together with the structural similarity of the active site

of *Tbb*MCA2 with that of *Tco*MCA5 and *Tvi*MCA5 suggests that the docking results can be compared across these three proteins.

Interactions between the ligands and the protein 'receptor' were determined using the results of the docking studies. This can be optimised when used in combination with molecular dynamic studies, whereby different conformational changes of the protein 'receptor' when binding to a ligand are modelled, as was reported for the identification of the interactions between the clathrin 'receptor' and its natural product inhibitor, bolinaquinone (Abdel-Hamid and McCluskey, 2014). A comparison of the number of interactions between the residues of the active site and the various substrates, inhibitors and ligands indicated that the commercial library substrates and inhibitors formed more interactions with the MCA 'receptor' than those of the Berg library ligands. For further optimisation of the Berg ligands and improvement of the IC_{50} values, the results of the commercial ligands need be taken into consideration. Both the Berg and commercial ligands displayed Pi-sulfur interactions with the secondary catalytic Cys and catalytic Cys in both *Tbb*MCA2 (Cys92 and Cys213) and the MCA5s (Cys81 and Cys202) which gives credibility to the possibility of a secondary catalytic Cys. This could explain why when both the catalytic dyad residues of *Tco*MCA5 were mutated, activity was still evident (Chapter 3). The Cys adjacent to the catalytic Cys (*Tbb*MCA2: Cys212, MCA5: Cys201) did not seem to interact with any residues of the various ligands.

The generation of a pharmacophore summarises the steric and electronic features required for optimal interaction of a ligand with a protein 'receptor' (Sliwoski *et al.*, 2013). Using the spatial positions of the commercial and Berg ligands which had the highest CDocker interaction energies, pharmacophore models were produced, with hydrogen bond acceptors and hydrophobic interactions being the main features of the interactions with the MCAs followed by ionisation and hydrogen bond donors. These features can be utilised for virtual screening of compound libraries for the identification of novel ligands with high binding affinities for the MCA 'receptor' (Sliwoski *et al.*, 2013). The use of molecular docking allowed for the selection of compounds which possessed higher binding affinities for cyclooxygenase-2 (COX-2) than the classical inhibitor, celecoxib (Puratchikody *et al.*, 2016). In addition, pharmacophore models can be used to guide the design of new ligands for the optimisation of the IC_{50} values (Yang, 2010; Sliwoski *et al.*, 2013).

Various kinetoplastid MCAs have been shown to play a role in the cell cycle (Helms *et al.*, 2006; Ambit *et al.*, 2008; Proto *et al.*, 2011; Laverrière *et al.*, 2012). As such,

inhibitors against these peptidases are valuable contributions towards a drug which targets the MCAs. The successful molecular docking of the Berg ligands in the active site of the 3D structure of *Tbb*MCA2 and homology modelled *Tco*MCA5 and *Tvi*MCA5 suggest that the phenotypic assay results (Berg *et al.*, 2012) were a reflection of the inhibition of MCAs. Together with that of the Berg library, the molecular docking of the commercial library provided insight into the interactions between the ligands and the protein 'receptor'. Combined with this information and the pharmacophore model, ligand optimisation and virtual screening can be implemented to identify a ligand with a high binding affinity for the MCAs. Since the inhibitors, substrates and ligands bound to both the single- (MCA5) and multicopy (MCA2) MCAs, and in different species, *T. b. brucei, T. congolense* and *T. vivax*, it is possible that these ligands are pan specific.

In the next chapter, work is presented on the application of phage display technology for the production of recombinant antibodies against OPB as a model antigen. This paves the way for applying this technology for the production of MCA specific antibodies for use in diagnostics as well as for parasite imaging in cellular research.

CHAPTER 5

PHAGE DISPLAY AND ANTIGEN DETECTION ELISA OF OLIGOPEPTIDASE B FROM *T. CONGOLENSE* AND *T. VIVAX*

5.1 Introduction

Due to the location of animal African trypanosomosis (AAT) infections in resource-poor, rural settings together with the waves of parasitaemia caused by antigenic variation, the development of an accurate, cost effective and easy to use diagnostic assay is imperative to the control of the spread of infections through the identification and treatment of infected individuals (Aksoy *et al.*, 2017; Diall *et al.*, 2017). One such diagnostic method is the immunochromatographic dipstick test, also known as a rapid diagnostic test (RDT), which has been successfully used for the detection of *T. congolense* and *T. vivax* (Boulangé *et al.*, 2017), *T. cruzi* (Luquetti *et al.*, 2003; Cardinal *et al.*, 2006) and *Leishmania infantum* (Reithinger *et al.*, 2002) infections.

An attractive antigen to be incorporated into an RDT is one which is released by the parasites into the host circulatory system and is stable. As anti-trypanosomal antibodies persist long after successful treatment (Paquet *et al.*, 1992; Yadav *et al.*, 2012), an antigen detection format would be able to differentiate between current and past infections. As such, parasite proteins which are present within the bloodstream of the infected host are attractive diagnostic antigens (Eyford *et al.*, 2013). An antigen detection ELISA was developed using species-specific monoclonal antibodies against the antigens released from the lysis of *in vitro* cultured *T. congolense* parasites (Nantulya *et al.*, 1987). This test was shown to be more than four times more sensitive than microhematocrit centrifugation technique in monitoring experimental *T. congolense* infections in goats and cattle (Masake and Nantulya, 1991).

One such potential diagnostic antigen is the serine peptidase, oligopeptidase B (OPB) that was used as a model antigen in the present study. It has been demonstrated that *Tbb*OPB (Morty *et al.*, 2005b), *Tco*OPB (Pinto Dias, 2006) and *Tev*OPB (Morty *et al.*, 2005a) are released from lysed, dead or dying parasites into the bloodstream, where they remain active (Morty *et al.*, 2005a; Munday *et al.*, 2011). The action of OPB has been implicated in the disruption of the hormone levels of the infected host (Morty *et al.*, 1999; Morty *et al.*, 2001; Morty *et al.*, 2005a) through the cleavage of peptide hormones in the blood of *T. b. brucei* infected rats (Tetaert *et al.*, 1993) and the reduced levels of regulatory peptides such as atrial natriuretic factor. This renders OPB a

desirable target for antigen detection. Anti-OPB antibodies were identified in sera from humans infected with *T. cruzi* and *L. infantum*, and *T. cruzi* experimentally infected rabbits using an antibody detection ELISA in an indirect format (Fernandes *et al.*, 2005). However, a previous study to detect anti-OPB antibodies in sera from *T. congolense* infected and non-infected cattle, in an antibody detection ELISA, was unsuccessful (Eyssen, 2013).

Thus, in order to develop an antigen detection ELISA assay, antigen specific antibodies are required. The established method of the production of polyclonal antibodies is by animal immunisation (Schirrmann *et al.*, 2011). Antibodies against the full length *Tco*OPB and *Tvi*OPB and peptides of *Tco*OPB have been produced in chickens and rabbits previously (Huson, 2006; Kangethe, 2011). However, *Tvi*OPB was poorly immunogenic in chickens, resulting in low chicken anti-*Tvi*OPB IgY titres (Eyssen, 2013). As such, the production of alternative antibodies was investigated. Monoclonal antibodies created by hybridoma technology (Köhler and Milstein, 1975), at a high cost, have limited application in the generation of human therapeutic antibodies and antibodies against toxic or highly conserved antigens (Winter and Milstein, 1991). *In vitro* technologies have an advantage over animal immunisation in that it is completely independent of the immune system and it comprises an unlimited source of binders, of which the DNA sequence is known (Schirrmann *et al.*, 2011). As a result, there would be no deviations between different batches as experienced with animal serum (Schirrmann *et al.*, 2011). One such *in vitro* technology is that of phage display which expresses antibodies, derived from combinatorial libraries of immunoglobulin V_H and V_L genes, on phage particles (Pini and Bracci, 2000).

Using a camelid library comprised of the variable heavy domain of camelid antibodies, V_HH, panning against the variable surface glycoprotein (VSG) of *T. b. brucei* resulted in a V_HH antibody which showed superior penetration within the spaces between VSG dimers, which is not accessible to the IgG and IgM antibodies produced during infection (Stijlemans *et al.*, 2004). This demonstrates that the small antibody fragments are able to penetrate between the narrow cavities of immune-evasive antigens as demonstrated in the improved the pharmacokinetics of tissue penetration during antibody-based cancer therapy (Chowdhury *et al.*, 1998; Deckert, 2009).

The naïve, semi-synthetic and synthetic libraries, also known as ‘single pot’ libraries, are able to identify antibody fragments, which in theory, bind to every possible antigen (Hust and Dübel, 2004). These single pot libraries serve as the molecular repertoire for

the process of phage display (Sblattero and Bradbury, 2000). The naïve *Nkuku*® library is not confined to any specific target and, as a result, facilitates the generation of monovalent scFv antibodies against an almost unlimited range of antigens (van Wyngaardt *et al.*, 2004). The scFv antibody (V_L and V_H) expressed using phage display, is able to retain the specific antigen binding affinity to that of the parent IgG.

The *Nkuku*® library was constructed from the rearranged immunoglobulin V genes from B cells (IgM) of non-immunised donor chickens (van Wyngaardt *et al.*, 2004). These genes were then ligated into the pHEN1 phagemid plasmid and inserted into bacteriophages to produce a phagemid (Fig. 1.8, Step A). Those phagemids which possess complementary determining regions (CDRs) to that of the immobilised target antigen are retained (Step F), used to infect *E. coli* cells and are rescued by helper phages (Step D). The resulting phagemids are isolated and panned against the immobilised antigen to select for scFvs with a high binding affinity for the antigen. The resulting scFv antibodies are myc-tagged, and may be produced as a minor phage coat protein (pIII) fusion proteins, scFv::pIII, if expressed in amber suppressor *E. coli* cells (Appendix D1). The panning process is illustrated in Fig. 5.1 and numbered as detailed in Fig. 1.8, with each step detailed in the methods section.

The production of phage displayed scFv antibodies against the OPB antigen from both *T. congolense* and *T. vivax*, and their application in an antigen detection ELISA towards the development of an AT diagnostic test is reported here. In addition, homology modelling of the scFv and its docking onto the OPB antigen is reported.

5.2 Materials and methods

5.2.1 Materials

Phage display library: The *Nkuku*® phage display library, M13KO7 helper phages and the amber suppressor TG1 {F' [*traD36 proAB*$^+$ *lacI*^{q} *lacZΔM15*]*supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5, (*$r_K^-m_K^-$*)*} *E. coli* cells, were obtained from Wouter van Wyngaardt and Jeanni Fehrsen (Immunology section, Onderstepoort Veterinary Institute).

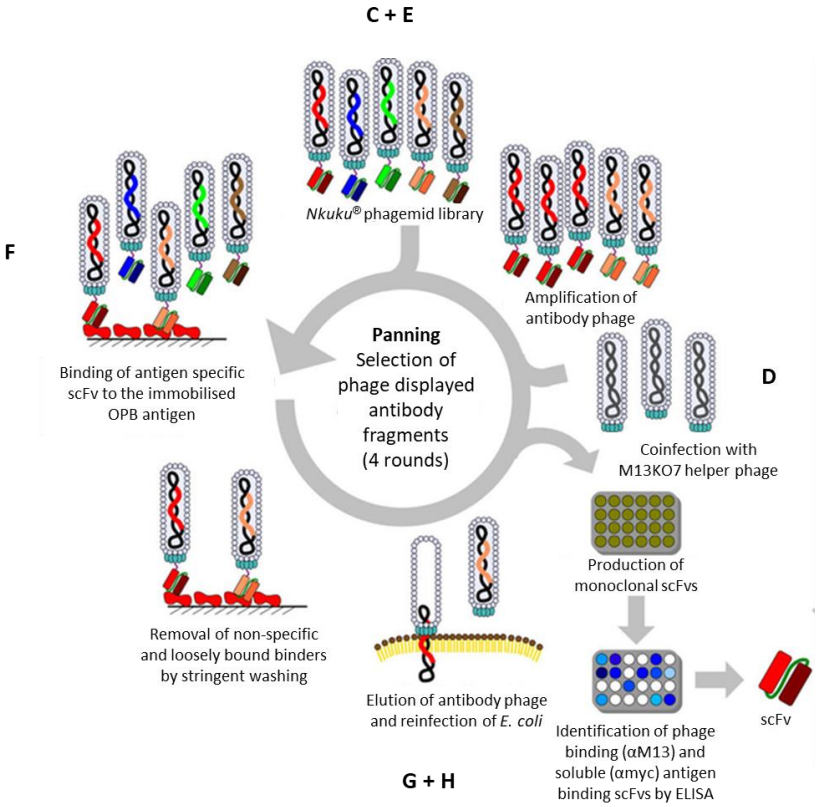


Figure 5.1: Selection of antigen specific scFv antibodies using panning. The image was adapted from Kuhn *et al*. (2016). The *Nkuku*® phagemid transfected *E. coli* cells, which display the scFv::pIII fusion protein are panned against the immobilised OPB antigen. The binders which are not OPB-specific, or bind weakly, are removed with stringent washing. The bound antibody phages are eluted, reinfected into the *E. coli* host, and coinfected with the M13KO7 helper phage. After amplification of the selected phagemid population, it is panned against the OPB antigen once more to facilitate the selection of highly specific antigen binders. This process was performed four times in total in this study. After the fourth round of panning, colonies of the reinfected *E. coli* cells containing the antibody phage specific to the OPB antigen, are expressed in 96-well culture plates. These clones are analysed by ELISA using antibodies against the phage coat protein (pIII) and the myc affinity tag for the identification of the scFv::pIII fusion proteins and soluble scFv proteins, respectively. The letters A-H correspond to those in Fig. 1.8 and each step is detailed in the methods section.

Biopanning: Nunc-Maxisorp 4 ml immunotubes (Catalogue number: 444474) were obtained from Nunc Intermed (Roskilde, Denmark), ELITE skim milk powder, for panning, from Clover (SA), bovine serum albumin (Catalogue number: A7906) from Sigma (St. Louis, MO. USA.), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (IPTG) from Fermentas (Vilnius, Lithuania), AluI and the 100 bp DNA ladder from New England Biolabs (Ipswich, MA, USA), ampicillin sodium salt from USB Corporation (Cleveland, OH, USA), kanamycin from Gibco, (Paisley, UK), and non-fat powdered milk for western blots and ELISAs, from Amresco (Solon, OH, USA). Bacteriological agar, tryptone and yeast extract were purchased from Merck Biolab (Darmstadt,

Germany). Buffer salts and other common chemicals were purchased from Merck (Germany) and Sigma (St. Louis, MO, USA) and were of the highest purity available.

Purification and quantification of phage displayed scFv: As per Section 2.2.1 and 3.2.1. The *Tco*OPB-Aminolink®, *Tvi*OPB-Aminolink® (Huson, 2006) and chicken anti-myc IgY hydrazide (Krause, 2016) resins were previously prepared.

Antibodies: Mouse anti-M13 IgG was purchased from Pierce (Rockford, IL, USA), goat anti-mouse IgG HRPO conjugate from Jackson ImmunoResearch (PA, USA) and rabbit anti-chicken IgY HRPO conjugate from Sigma (St. Louis, MO. USA). The affinity purified chicken anti-myc IgY, raised against the EQKLISEEDL peptide, was produced previously by Rob Krause (2016). Both the rabbit anti-*Tco*OPB IgG and the affinity purified chicken anti-peptide 4 IgY were produced previously by Laura Huson (2006). The peptide 4 corresponds to the 282 to 299 amino acid residues of the *Tco*OPB protein, VRKREKNVRYEVEMHGT.

3D structures of scFv antibodies: The 3D structures, solved by X-ray diffraction, of scFv complexed to human gankyrin (PDB: 4NIK, 2.5 Å) (Robin *et al.*, 2014), cross reactive antibody which binds to respiratory syncytial virus and human metapneumovirus (5U68, 3.08 Å) (Wen *et al.*, 2017), the scFv against prostate specific antigen (4P49, 1.4 Å) (Conroy *et al.*, 2014), the scFv against cardiac troponin I (4P48, 1.35 Å) (Conroy *et al.*, 2014), the scFv against the major allergen from *Blatella germanica* (4OUO, 1.8 Å) (Mueller *et al.*, 2014) and *Tbb*OPB (4BP9, 2.8 Å) (Canning *et al.*, 2013) were obtained from the RCSB Protein Data Bank (www.rcsb.org)(Berman *et al.*, 2000).

Discovery Studio docking software: The Centre for High Performance Computing (Cape Town, South Africa) is acknowledged for the provision of the computational resources necessary to conduct the prediction of the scFv interaction with the target *Tco*OPB antigen.

Sera: Sera from infected and non-infected cattle were obtained from ClinVet International (PTY) LTD, Bloemfontein, South Africa.

5.2.2 Preparation of TG1 *E. coli* stock for log-phase culture

Escherichia coli TG1 cells were streaked onto TYE plates [1% (w/v) tryptone, 0.5% yeast extract, 0.8% (w/v) NaCl, 1.5% (w/v) agar] without antibiotic, and grown at 37 °C

for 16 h. A single colony was inoculated into 2xYT (5 ml), without antibiotic, and grown at 37 °C for 16 h at 220 rpm. A 1:100 dilution of the overnight culture was performed in fresh 2xYT, without antibiotic, and incubated at 37 °C until the OD_{600} reached 0.5. This log-phase culture was stored at 4 °C for a maximum of 7 days.

5.2.3 Expression and titration of M13KO7 helper phages

For the titre determination of the M13KO7 helper phages, a serial dilution of M13KO7 in PBS (100 µl) were incubated with log-phase TG1 *E. coli* cells (100 µl) at RT for 5 min. To each dilution, pre-warmed 2xYT [containing 0.7% (w/v) agar] without antibiotic, at 42 °C, was added and immediately plated onto 2xYT plates, without antibiotic and incubated at 37 °C for 16 h.

A single kanamycin resistant (Kan^r) M13KO7 plaque was incubated with log-phase TG1 *E. coli* cells (40 µl) in 2xYT (4 ml), without antibiotic, at 37 °C for 1 h at 100 rpm. Fresh 2xYT (800 ml), without antibiotic, was added and incubated for an additional hour. Thereafter, kanamycin (50 µg/ml) was added and incubated at 37 °C for 16 h at 200 rpm. The cells were pelleted (10 800*g*, 5 min, 4 °C) and to the phage containing supernatant, 1/4 volume of 20% (w/v) PEG 6000 in 2.5 M NaCl added. The mixture was incubated on ice for 30 min before the phages were pelleted (10 800*g*, 15 min, 4 °C). After resuspension in PBS (6 ml) and filtering through a 0.22 µM filter, the titre of the M13KO7 helper phages was determined as above. The M13KO7 helper phages were stored at a concentration of $1x10^{12}$ pfu/ml in 15% (v/v) glycerol at -80 °C.

5.2.4 Expression and titration of the *Nkuku*® phagemid library

The *Nkuku*® phagemid library (Fig. 1.8, Step A) was provided in *E. coli* cells (step B). Following Step C, the ampicillin resistant (Amp^r) *Nkuku*® library glycerol stock was added to 2xYT [containing 100 µg/ml ampicillin, 2% (w/v) glucose, 100 ml], until an initial OD_{600} of 0.05 was obtained. Incubation at 37 °C at 240 rpm was carried out until an OD_{600} of 0.5 was obtained.

Following Steps D and E, to the *Nkuku*® culture (100 ml), $8x10^9$ pfu/ml M13KO7 helper phages were added and incubated at 37 °C for 30 min without agitation, followed by a further 30 min at 100 rpm. The cells were pelleted (3 300*g,* 10 min, 4 °C), resuspended in 2xYT (containing 100 µg/ml ampicillin, 25 µg/ml kanamycin, 2 L), and incubated at 30 °C for 16 h at 240 rpm.

The cells were pelleted (3 300*g,* 10 min, 4 °C), and to the phagemid containing supernatant,1/4 volume of 20% (w/v) PEG 6000 in 2.5 M NaCl added (250 ml). After 1 h on ice, the phagemids were pelleted (3 300 *g*, 15 min, 4 °C), resuspended in PBS (20 ml), and pelleted once again (11 000 *g*, 2 min, 4 °C). The phagemids were filtered through a 0.22 µM filter and used for the first round of panning.

5.2.5 Panning of the *Nkuku*® library against oligopeptidase B from *T. congolense* and *T. vivax*

Panning was performed as previously described by van Wyngaardt *et al.* (2004) as shown in Fig. 1.8 and Fig. 5.1, Step F. Briefly, immunotubes were coated with 100, 50, 10 and 1 µg/ml antigen (3.5 ml) diluted in PBS, for pans 1, 2, 3 and 4, respectively, for 16 h at 4 °C. The immunotubes were washed three times with PBS and blocked with either 2% skim milk powder-PBS (MP) or 3% (w/v) BSA-PBS (BP), two of the most commonly used blocking agents used in panning, at RT for 1 h. The immunotubes were then washed twice with 0.1% (v/v) Tween-20-PBS (PBS-T) followed by PBS. During the blocking of the immunotube, *Nkuku*® phagemid was diluted to 10^{12} pfu/ml in MP or BP [containing 0.1% (v/v) Tween-20, 3.5 ml] and incubated at 37 °C for 30 min. This blocking step assists in the reduction of binding to unrelated peptides (Thomas *et al.*, 2010; Vodnik *et al.*, 2011).

The MP and BP treated *Nkuku*® phagemid was added to the blocked immunotube after washing and incubated at RT for 30 min with gentle agitation followed by 1.5 h without agitation. The immunotube was washed twenty times with PBS-T and a further twenty times with PBS. Following Fig. 5.1, step G, freshly prepared 100 mM triethylamine (1 ml) was added to the immunotube and mixed with agitation using an end-over-end rotator at RT for 10 min. The eluted phage was neutralised by the addition of 1 M Tris-HCl buffer, pH 7.4 (0.5 ml). As per step H in Fig. 5.1, log-phase TG1 *E. coli* cells (5 ml) were transfected with the neutralised phages (1 ml) in a water bath at 37 °C for 30 min. The cells were pelleted (3 300*g*, 10 min, 4 °C), resuspended in 2xYT (1 ml), without antibiotic, plated onto three TYE plates [containing 100 µg/ml ampicillin, 2% (w/v) glucose] and incubated at 30 °C for 16 h. To determine whether or not the TG1 *E. coli* cells had been transfected previously, log-phase TG1 *E. coli* cells were also plated onto TYE plates [containing 100 µg/ml ampicillin, 2% (w/v) glucose].

Instead of using *Nkuku*® phagemid, all subsequent panning rounds will be screened using the resulting phagemids from the previous round of selection against OPB as per

step C in Fig. 5.1. The resulting colonies were resuspended in 2xYT (8 ml), without antibiotic, using a “hockey stick” and stored in 15% glycerol at -80 °C. This is the glycerol stock from the first round of panning. For amplification of the phages, the glycerol stock from the first panning round was added to 2xYT [containing 100 µg/ml ampicillin, 2% (w/v) glucose, 50 ml] to achieve an initial OD_{600} of 0.05. This was then incubated at 37 °C at 220 rpm until an OD_{600} of 0.5 was obtained. To this culture, $8x10^9$ pfu/ml M13KO7 helper phages were added and incubated at 37 °C for 30 min without agitation as per step D and E in Fig. 5.1. The cells were pelleted (3 300 *g,* 10 min, 4 °C), resuspended in 2xYT [containing 100 µg/ml ampicillin, 25 µg/ml kanamycin, 25 ml], and incubated at 30°C for 16 h at 220 rpm.

The cells were pelleted (3 300*g,* 10 min, 4 °C) and to the phage containing supernatant, 1/4 volume of 20% (w/v) PEG 6000 in 2.5 M NaCl was added (5 ml). After 1.5 h at 4 °C, the phages were pelleted (3 300 *g*, 15 min, 4 °C), resuspended in PBS (1 ml), and pelleted once again (11 000 *g*, 2 min, 4 °C). The phages were filtered through a 0.22 µM filter and used for the next round of panning as per step F in Fig. 5.1. This process was repeated for a total of 4 rounds with each immunotube being coated with decreasing concentrations of antigen, which functions to enrich for specific antigen binders.

5.2.6 Determination of enrichment of panned phages

Purified recombinant OPB (Eyssen, 2013), diluted in PBS (1 µg/ml, 100 µl/well), was used to coat the wells of 96-well Nunc-Immuno™ Maxisorp ELISA plates for 16 h at 4 °C. The coating solution was discarded, and the plates were blocked with blocking buffer (MP and BP, 200 µl/well) and incubated at 37 °C for 1 h. Phages from pans one to four, including the *Nkuku*® phagemid [1:10 in MP or BP, each containing 0.1% (v/v) Tween-20, 100 µl/well] were added and incubated at 37 °C for 2 h. The wells were washed three times with PBS-T using a BIOTEK® ELx50™ Microplate washer after which the mouse anti-M13 IgG [1:8 000 in MP or BP, each containing 0.1% (v/v) Tween-20, 100 µl/well], was added and incubated at 37 °C for 1 h. The wells were washed as before, followed by the addition of the goat anti-mouse IgG HRPO conjugate [1: 1 000 in MP or BP, each containing 0.1% (v/v) Tween-20, 100 µl/well], and incubated at 37 °C for 1 h. The wells were washed as before and the ABTS·H_2O_2 chromogen substrate solution [0.05% (w/v) ABTS, 0.0015% (v/v) H_2O_2 in 0.15 M citrate-phosphate buffer, pH 5.0, 100 µl/well] added. The plate was incubated in the

dark prior to measuring the absorbance at 405 nm using a FLUORStar Optima spectrophotometer every 15 min for 1 h.

5.2.7 Selection of phage displayed scFv clones

The glycerol stocks from the selected enriched pans were serially diluted on TYE plates [containing 100 µg/ml ampicillin, 2% (w/v) glucose] at 30 °C for 16 h, in order to obtain individual colonies (30 to 300). Forty eight individual colonies for each antigen were selected and inoculated into a 96-well plate containing 2xYT (100 µl/well), without antibiotic. After incubation at 37 °C for 16 h at 220 rpm, a final concentration of 20% (v/v) glycerol was added to each well and stored at -80 °C. This served as the master plate.

In a new 96-well plate, an aliquot from each well of the master plate (5 µl) was used to inoculate 2xYT [containing 100 µg/ml ampicillin, 2% (w/v) glucose, 150 µl/well], and incubated at 37 °C for 2.5 h at 220 rpm. To each well, $2x10^9$ pfu/ml M13KO7 was added and incubated for a further 30 min without agitation. The cells were pelleted (600*g*, 10 min, 4 °C), resuspended in 2xYT (containing 100 µg/ml ampicillin, 25 µg/ml kanamycin, 150 µl/well) and incubated at 30 °C for 16 h at 220 rpm. The cells were pelleted and the phage containing supernatant (50 µl) was added to a blocked ELISA plate.

The ELISA was performed as described in Section 5.2.6, with only MP being used as the blocking and the dilution agent. Instead of using the isolated phagemids, the phage containing expression supernatants (50 µl/well) were used and diluted in 2xMP [containing, 0.1% (v/v) Tween-20, 50 µl/well].

5.2.8 Selection clones expressing soluble scFv

In a 96-well plate, an aliquot from each well of the master plate (5 µl) was used to inoculate 2xYT [containing 100 µg/ml Amp, 2% (w/v) glucose, 150 µl/well] and incubated at 30 °C for 16 h at 220rpm. To the overnight culture (5 µl), 2xYT [containing 100 µg/ml ampicillin, 0.1% (v/v) glucose, 100 µl/well] was added and incubated at 37 °C for 2.5 h at 220rpm. After the addition of 2xYT (containing 100 µg/ml ampicillin, 3 mM IPTG, 50 µl/well), the plate was incubated at 30 °C for 16 h at 220rpm. The cells were pelleted (600*g*, 10 min, 4 °C) and the phage containing supernatant (50 µl) was added to a blocked ELISA plate.

The ELISA was performed as described in Section 5.2.6, with only MP being used as the blocking and dilution agent, and the dilution of phage containing supernatants (50 µl/well) in 2xMP [containing, 0.1% (v/v) Tween-20, 50 µl/well]. The bound phages were detected using affinity purified chicken anti-myc IgY (0.5 µg/ml, 100 µl/well) followed by rabbit anti-chicken IgY HRPO conjugate (1:5 000, 100 µl/well).

5.2.9 Characterisation of OPB-specific scFv producing clones

Clones which were identified as OPB-specific in both the phage displayed scFv (Section 5.2.7) and expressed, soluble scFv (Section 5.2.8) ELISAs, were selected and grown in 2xYT [containing 100 µg/ml Amp, 2% (w/v) glucose, 5 ml] at 37 °C for 16 h with agitation. The plasmid DNA of the recombinant colonies was isolated using the GeneJet™ Plasmid Miniprep Kit according to the manufacturer's instructions. The isolated plasmid DNA was used for PCR amplification, DNA fingerprinting to determine the scFv insert diversity between the selected clones, and for sequencing.

The final concentrations of the PCR master mix were: 0.25 µM of the OP52 forward (5' CCC TCA TAG TTA GCG TAA CG 3') and M13 reverse (5' CAG GAA ACA GCT ATG AC 3') primers, 1xFIREpol® Taq polymerase reaction buffer, 2.5 mM $MgCl_2$, 1 U FIREpol® Taq polymerase and 0.25 mM dNTPs in a total reaction volume of 20 µl. The PCR amplification of the recombinant scFv gene, was performed with incubation at 95 °C for 2 min as the initial DNA denaturation step, followed by 40 cycles of 95 °C for 10 s, 45 °C for 15 s and 72°C for 1 min. A final elongation step was carried out at 72 °C for 7 min. The colony PCR products (10 µl) and samples of the plasmid DNA (5 µl) were electrophoresed on a 1.5% (w/v) agarose gel, containing ethidium bromide (0.5 µg/ml), in 1xTAE buffer.

The high frequency endonuclease, AluI, was used to determine the diversity of the scFv between the individual clones. Due to its notable ability to differentiate between highly related antibody sequences, AluI was used instead BstOI (Hawlisch *et al.*, 2000). Restriction digestion with 1 U of AluI was performed on plasmid DNA (5 µl) in 10xcutsmart buffer at 37 °C for 15 min. Samples of the digestion (5 µl) were electrophoresed on a 3% (w/v) agarose gel, containing ethidium bromide (0.5 µg/ml), in 1xTAE buffer.

The plasmid DNA of the various PCR clones were sequenced at the Central Analytical Facility, Stellenbosch University, South Africa, using the OP52 and M13 forward and reverse primers, respectively.

5.2.10 3D modelling of *Tco*OPB and scFv and their interactions

In a similar way as outlined in Section 4.2.2, suitable templates to model scFv and *Tco*OPB were obtained using SwissModel (Biasini *et al.*, 2014).

One of the computational prediction tools, EpiPred (Krawczyk *et al.*, 2014), was accessed from the structural antibody prediction server (SAbPred) (Dunbar *et al.*, 2016) developed by the Oxford Protein Informatics Group (http://opig.stats.ox.ac.uk/webapps/sabdab-sabpred/EpiPred.php). The generated epitope predictions can be used to improve the performance of antibody-antigen docking. Using the V_H and V_L sequences of the OPB-specific scFv, together with the homology modelled *Tco*OPB antigen, EpiPred was used for the prediction of the structural epitopes of *Tco*OPB specific to the scFv.

Antibody-antigen docking was used to predict the interaction of the OPB-specific scFv onto *Tco*OPB using the ZDock protocol from Discovery Studio (Dassault Systèmes BIOVIA).

5.2.11 Expression and purification of the scFv antibodies

Single colonies of the selected scFv clones were used to inoculate 2xYT (5 ml), without antibiotic, and incubated at 37 °C for 16 h at 220rpm. The overnight culture (1 ml) was diluted 1:100 in 2xYT [containing 100 µg/ml Amp, 2% (w/v) glucose], without antibiotic, and incubated at 37 °C at 220rpm until an OD_{600} of 0.9 was obtained. The cells were pelleted (3 000*g*, 10 min, 4 °C) and resuspended in terrific broth (containing 100 µg/ml ampicillin, 1 mM IPTG, 100 ml), and incubated at 30 °C for 16 h at 220 rpm. The cells were pelleted (4 000*g*, 10 min, 4 °C) and the supernatant retained. The cells were resuspended in 1/4 volume of ice cold sucrose buffer [50 mM Tris-HCl buffer, pH 7.4, 1 mM EDTA, 20% (w/v) sucrose] and incubated on ice for 10 min with gentle agitation. After pelleting the cells (4 000*g*, 10 min, 4 °C), the periplasmic fraction was retained, and the cells were resuspended in 1/10 volume ice cold 5 mM $MgCl_2$. Following incubation on ice for 10 min with gentle agitation, the cells were pelleted (8 000*g*, 10 min, 4 °C) and the resulting osmotic shock fraction was retained. Samples of the periplasm (P) and osmotic shock (OS) fluid were analysed by a 12.5% reducing SDS-PAGE gel (Laemmli, 1970), transferred onto nitrocellulose, as per Section 2.2.4. The blot was probed with affinity purified chicken anti-myc IgY (0.5 µg/ml) followed by rabbit anti-chicken IgY HRPO conjugate (1:5 000).

5.2.12 Affinity purification of expressed scFv

Purification of scFv using the *Tco*OPB- and *Tvi*OPB-affinity matrices as well as the chicken anti-myc affinity matrix was performed as outlined in Section 2.2.7.2 using the scFv containing osmotic shock fraction (8 ml). Samples of each fraction were analysed by 12.5% reducing SDS-PAGE gels (Laemmli, 1970), stained with Coomassie Blue R-250 and transferred onto nitrocellulose, as detailed in Section 2.2.4. The blot was probed with affinity purified chicken anti-myc IgY (0.5 µg/ml) followed by rabbit anti-chicken IgY HRPO conjugate (1:5 000).

5.2.13 Antigen detection of recombinant and native oligopeptidase B by scFv

The ability of the purified scFv to detect both the recombinant and the native *Tco*OPB present in the digitonin fractionated *T. congolense* parasites (Section 3.2.5) was determined using a dot blot. The primary antibodies used were rabbit anti-*Tco*OPB IgG (10µg/ml), affinity purified chicken anti-peptide 4 IgY (5 µg/ml), both made previously by Laura Huson (2006), chicken anti-*Tco*OPB IgY (10 µg/ml) (Eyssen, 2013) and purified scFv (1:25). The use of scFv required a bridging antibody of affinity purified chicken anti-myc IgY (0.5 µg/ml) before the addition of the secondary detection antibodies, rabbit anti-chicken IgY HRPO conjugate (1:5 000) and goat anti-rabbit IgG HRPO conjugate (1: 2500). Detection was achieved using the ECL western blotting substrate.

An antigen detection ELISA was performed using the scFv as a capture antibody and either the chicken or rabbit anti-*Tco*OPB antibodies as the detection antibody. Purified scFv (1: 100) was used to coat the wells of 96-well Nunc-Immuno™ Maxisorp ELISA plates (100 µl/well) for 16 h at 4 °C. The coating solution was discarded, and the plates were blocked with blocking buffer [0.5% (w/v) BSA-PBS, 0.1% (v/v) Tween-20, 200 µl/well] and incubated at 37 °C for 1 h. Sera from *T. congolense* infected and non-infected cattle diluted in blocking buffer (1:100, 100 µl/well) were added and incubated at 37 °C for 1 h. The wells were washed and either the rabbit anti-*Tco*OPB IgG or the chicken anti-*Tco*OPB IgY antibody, diluted in blocking buffer (1 µg/ml, 100 µl/well), was added and incubated at 37 °C for 1 h. The wells were washed as before, and either the goat anti-rabbit IgG HRPO or the rabbit anti-chicken IgY HRPO conjugate, diluted in blocking buffer (1:15 000, 100 µl/well), was added and incubated at 37 °C for 1 h. The wells were washed and the ABTS·H_2O_2 chromogen substrate solution [0.05% (w/v) ABTS, 0.0015% (v/v) H_2O_2 in 0.15 M citrate-phosphate buffer,

pH 5.0, 100 µl/well] was added. The plate was incubated in the dark for 15 min prior to reading the absorbance at 405 nm using the FLUORStar Optima Spectrophotometer in 15 min intervals until absorbance values of above 1.0 were reached.

5.3 Results

5.3.1 Enrichment of scFv specific for oligopeptidase B from *T. congolense* and *T. vivax*

Using the *Nkuku*® library, scFv antibodies with complementary epitopes to OPB were selected for. Little to no enrichment was observed when using milk powder (MP) to block the immunotubes compared to that of BSA (BP) (Fig. 5.2). With the decrease in antigen coating concentration which was used as bait, together with stringent washing of non-specific binders during the four panning rounds, enrichment was observed. The *Nkuku*® phagemid as well as the BSA coating control displayed low absorbance values. Enrichment occurred at pan 3 and more so at pan 4 for *Tco*OPB. The only pan where enrichment was evident for *Tvi*OPB was at pan 4.

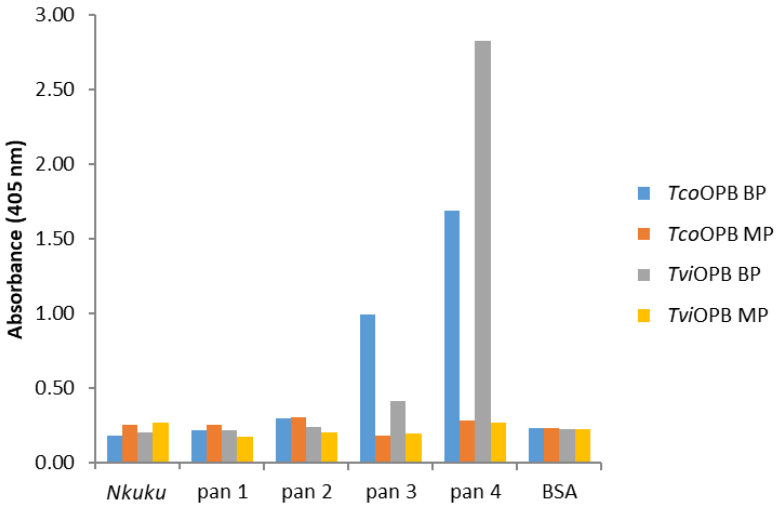


Figure 5.2: Enrichment ELISA of the isolated phages from each round of panning with oligopeptidase B from *T. congolense* and *T. vivax*. ELISA plates were coated with *Tco*OPB, *Tvi*OPB or BSA (1 µg/ml in PBS, pH 7.2), blocked with 3% (w/v) BSA-PBS and incubated with isolated phages (1:10) from each round of panning using immunotubes blocked with either 3% (w/v) BSA-PBS (BP) or 2% (w/v) skim milk powder-PBS (MP). Mouse anti-M13 IgG (1:8 000), followed by goat anti-mouse IgG HRPO (1:1 000) and ABTS·H_2O_2 were used as the detection system. The absorbance readings at 405 nm represent the average of duplicate experiments at 60 min development.

	Positive phage binders						Positive soluble scFv						Colony PCR						Sequenced						Positive by all					
*Tco*OPB pan 3	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
A	0.251	0.182	0.188	0.205	0.194	0.193	0.316	0.251	0.486	0.285	0.257	0.282																		
B	0.249	0.242	0.254	0.216	0.238	0.228	0.297	0.462	0.424	0.275	0.286	0.281																		
C	0.222	0.225	0.215	0.23	0.236	0.21	0.299	0.463	0.31	0.305	0.301	0.26																		
D	0.24	0.232	0.24	0.241	0.308	0.528	0.292	0.504	0.688	0.609	0.296	0.396		1000	1000	1000		1000												
E	0.284	0.212	0.213	0.498	0.582	0.726	0.567	0.277	0.726	0.618	0.744	0.424	1000		1000	1000	1000	1000												
F	0.294	0.288	0.231	0.227	0.268	0.227	0.424	0.543	0.325	0.302	0.332	0.313		1000																
G	0.24	0.256	0.325	0.225	0.242	0.372	0.329	0.479	0.435	0.353	0.338	0.417																		
H	0.264	0.276	0.244	0.253	0.334	0.285	0.354	0.355	0.817	0.343	0.548	0.705			1000		1000	1000												

*Tco*OPB pan 4	7	8	9	10	11	12	7	8	9	10	11	12	7	8	9	10	11	12	7	8	9	10	11	12	7	8	9	10	11	12
A	0.211	0.202	0.221	0.227	0.496	0.192	0.266	0.256	0.251	0.24	0.224	0.242					0													
B	0.219	0.226	0.481	0.22	0.221	0.371	0.271	0.271	0.391	0.267	0.255	0.349		500																
C	0.277	0.325	0.36	0.373	0.218	0.25	0.262	0.266	0.38	0.265	0.621	0.288					1000													
D	0.483	0.321	0.438	0.231	0.46	0.269	0.291	0.378	0.434	0.292	0.271	0.272	500	500																
E	0.379	0.239	0.218	0.223	0.756	0.286	0.268	0.285	0.276	0.274	0.327	0.29																		
F	0.426	0.375	0.371	0.428	0.297	0.363	0.302	0.32	0.311	0.319	0.311	0.614	500			500		1000												
G	0.249	0.223	0.215	0.256	0.245	0.504	0.614	0.324	0.31	0.62	0.314	0.335	1000			1000		500												
H	0.302	0.792	0.275	0.535	0.34	0.269	0.344	0.377	0.346	0.336	0.337	0.337		500		500														

*Tvi*OPB pan 4	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
A	0.224	0.322	0.199	0.324	0.202	0.187	0.368	0.368	0.265	0.217	0.213	0.404																		
B	0.304	0.229	0.285	0.499	0.251	0.338	0.398	0.278	0.318	0.248	0.428	0.395																		
C	0.231	0.386	0.21	0.212	0.201	0.198	0.867	0.533	0.278	0.653	0.491	0.646	1000	1000		1000		1000												
D	0.275	0.223	0.543	0.451	0.214	0.212	0.589	0.32	0.488	0.49	0.707	0.612	1000		1000	1000	1000	1000												
E	0.238	0.333	0.215	0.209	0.204	0.199	0.306	0.25	0.612	0.393	0.267	0.259			1000															
F	0.298	0.223	0.235	0.221	0.214	0.215	0.469	0.519	0.33	0.311	0.326	0.272		1000																
G	0.236	0.215	0.211	0.225	0.206	0.208	0.323	0.268	0.56	0.341	0.279	0.339			1000															

Figure 5.3: Results of various methods to select the best clones from *Tco*OPB and *Tvi*OPB pans 3 and 4 in TG1 *E. coli* cells for the expression of scFv antibodies. Using the absorbance values from the phage binding ELISA (green) together with those of the soluble scFv ELISA (pink), the plasmid DNA of the selected clones were subjected to PCR amplification of the scFv sequence and the resulting product sizes are indicated in bp (orange). The plasmid DNA of the positive clones were sequenced (yellow), and those which were positive using all four selection methods, are highlighted in blue.

5.3.2 Selection and characterisation of individual positive scFv clones

For pans three and four of *Tco*OPB and pan four of *Tvi*OPB, in TG1 *E. coli* cells, 48 single clones were randomly selected and tested in phage binding (Appendix D2) and soluble scFv ELISA formats (Appendix D3). In order to determine which of the clones were best for use in expression, the plasmid DNA of positive clones were subjected to colony PCR, to amplify the scFv gene, and DNA fingerprinting using AluI digestion (Appendix D4). The clones which had a high absorbance value in each of the ELISA formats and possessed a 1 000 bp amplified product were sequenced (Fig. 5.3).

Absorbance cut-off values for the two ELISA formats had to be assigned for the evaluation of the best clones (Appendix D5). Those clones which possessed values above the cut-off values and had a 1 000 bp amplified PCR product, were sequenced. An ideal scFv would be positive in both the phage binding and soluble ELISA formats (Rakabe, 2008). As such, clones E4 and E5 from *Tco*OPB pan three, and C2 from *Tvi*OPB pan four, all in TG1 *E. coli* cells, were selected for scFv expression (Fig. 5.3). These three clones will be referred to at E4, E5 and C2, respectively.

5.3.3 scFv sequencing, 3D modelling, and interaction with *Tco*OPB

5.3.3.1 scFv sequencing

Following the selection of positive clones, the sequences of the isolated plasmid DNA from both *Tco*OPB pans three and four, together with *Tvi*OPB at pan four, were determined and aligned with the immunoglobulin heavy (GenBank® BAE80128.1) and lambda (GenBank® BAB47352.1) chain sequences of *Gallus gallus* and referred to as the germline. Each of the sequenced clones possessed the exact same sequences in each of the complementary determining regions (CDRs) in both the V_H and V_L chains (Fig. 5.4). This may indicate a highly conserved region in OPB between the *T. congolense* and *T. vivax* species.

5.3.3.2 Homology modelling

A 81.3% sequence identity is shared between *Tco*OPB and *Tbb*OPB, the latter whose 3D structure has been determined by X-ray diffraction (Canning *et al.*, 2013). A homology model of *Tco*OPB was generated with a -Qmean of -0.83 using *Tbb*OPB as

the template. Slight variations were evident at some loops, but overall, the homology modelled *Tco*OPB overlaid well with that of *Tbb*OPB (Fig. 5.5). An enlarged image of the template and query sequence alignment is given in Appendix D6.

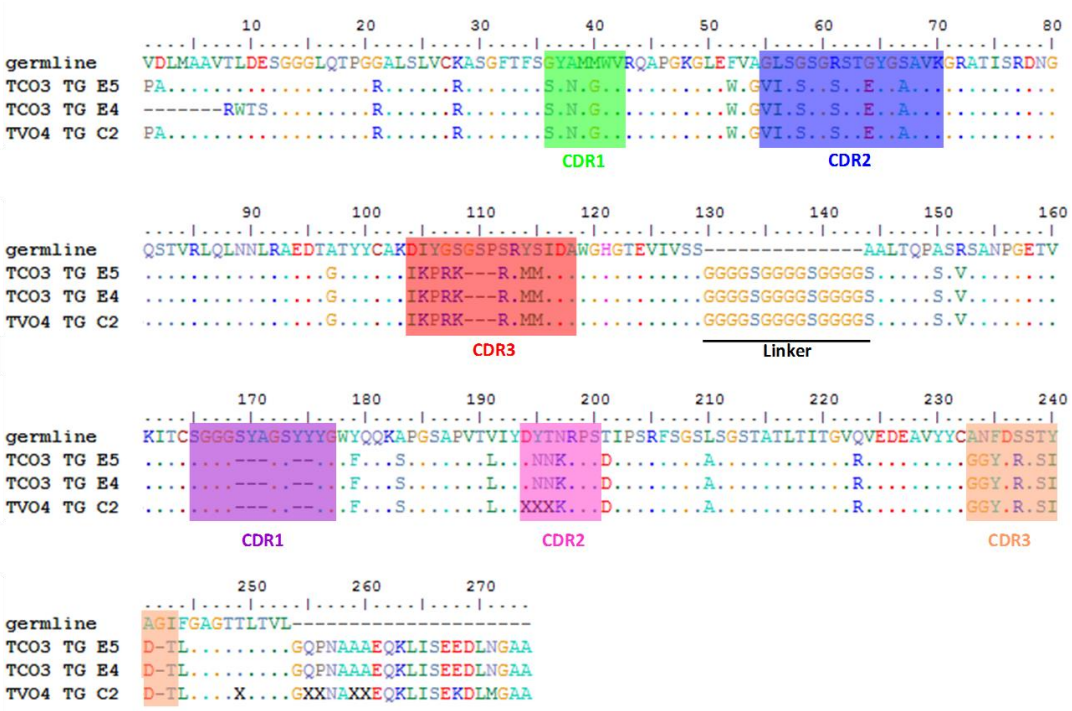


Figure 5.4: Multiple sequence alignment of scFv clones panned against *Tco*OPB and *Tvi*OPB. The sequences of clones E5 and E4 selected from *Tco*OPB pan three (TCO3) and clone C2 selected from *Tvi*OPB pan four (TVO4) were aligned against the germline (GenBank® BAE80128.1 and BAB47352.1) using the BioEdit software (Hall, 1999). The V_H CDR1, -2 and -3 regions are highlighted in green, blue and red, and the V_L CDR1, -2 and -3 regions in purple, pink and light orange, respectively.

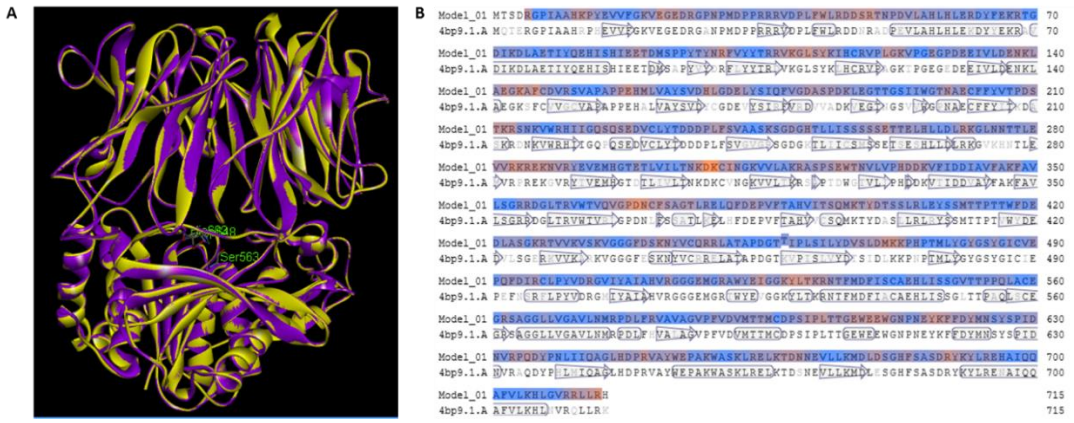


Figure 5.5: Homology modelling of *Tco*OPB. Using the 3D structure from the Protein Data Bank (Berman *et al.*, 2000) of *Tbb*OPB, 4BP9 (Canning *et al.*, 2013), as the template (yellow), the 3D structure of *Tco*OPB (purple) was **(A)** modelled using SwissModel (Biasini *et al.*, 2014) and superimposed. **(B)** The alignment of the OPB sequences is shown with the Qmean scores indicated in blue (high quality prediction) and red/orange (low quality prediction). (See Appendix D6 for an enlarged version of panel B).

To date, the 3D structures for only three chicken scFvs have been solved by X-ray diffraction, 4P48 (Conroy *et al.*, 2014), 4P49 (Conroy *et al.*, 2014) and 4OUO (Mueller *et al.*, 2014), in addition to scFvs from other phage libraries, 4NIK (Robin *et al.*, 2014) and 5U68 (Wen *et al.*, 2017). Results from the SwissModel software indicated that of the five scFvs, 4NIK, 5U68 and 4P49 shared the highest identity with the E5 scFv, 56.8, 58.3 and 54.3%, respectively. The E5 scFv was modelled using all three templates and are shown in Fig. 5.6. The lowest Qmean values were obtained with the 4NIK model (panel A) indicating a better model, followed by 5U68 (panel B) and 4P49 (panel C). Overall, the 4NIK and 5U68 models overlaid better on E5 scFv compared to that of 4P49, which only modelled the V_H region of the E5 scFv. As such, the E5 scFv model, built using the 4NIK template, was selected for the docking interaction studies with the homology modelled *Tco*OPB.

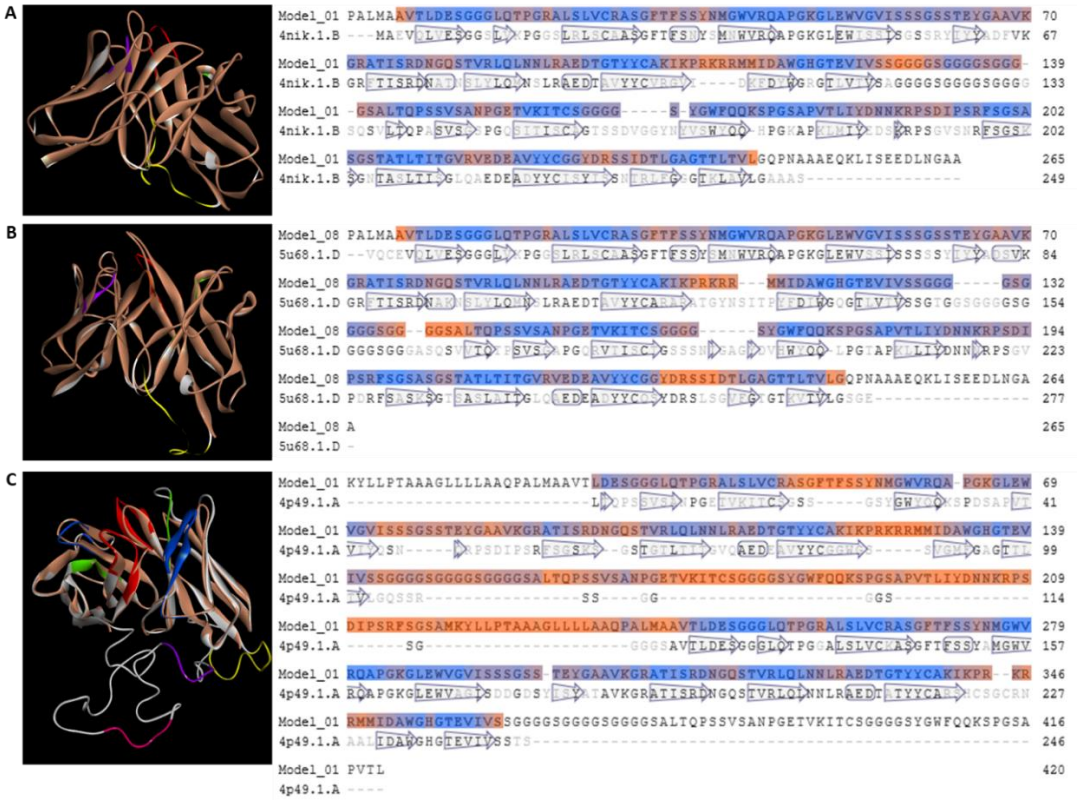


Figure 5.6: Homology modelling of scFv clone E5, from *Tco*OPB pan 3, using different scFv templates. The E5 anti-*Tco*OPB scFv antibody (white with coloured CDR regions) was modelled using and superimposed onto **(A)** 4NIK, scFv complexed to human gankyrin (Robin *et al.*, 2014), **(B)** 5U68, cross reactive antibody which binds to respiratory syncytial virus and human metapneumovirus (Wen *et al.*, 2017), and **(C)** 4P49, scFv against prostate specific antigen using the chicken anti-prostate specific antigen (Conroy *et al.*, 2014), all of which were coloured beige. The structures were obtained from the Protein Data Bank (Berman et al., 2000). The E5 scFv V_H CDR1, -2 and -3 are highlighted in green, blue, red, and those of the V_L in purple, pink and light orange, respectively. The linker region is highlighted in yellow. The alignment of the scFv sequences is shown together with the Qmean scores indicated in blue (high quality prediction) and red/orange (low quality prediction).

The validation of the generated homology modelled structures of both *Tco*OPB and E5 scFv, using the ProSA software and Ramachandran plots, are reported in Appendix D7.

5.3.3.3 scFv and *Tco*OPB interactions predictions

Using Predict7, an epitope prediction program (Cármenes *et al.*, 1989), seven immunogenic epitopes were identified in *Tco*OPB and anti-peptide antibodies produced in chickens (Huson, 2006). They corresponded to amino acids 20-37 (light green), 60-76 (red), 124-142 (purple), 282-299 (orange), 418-436 (pink), 658-671 (blue) and 686-698 (dark green) and were numbered 1 to 7, respectively (Fig. 5.7, panel A).

The EpiPred epitope predictions can be used to improve the performance of antibody-antigen docking. The hinge region is comprised of two polypeptide strands, residues 87-95 and 433-440 (Canning *et al.*, 2013), the first of which was not identified as an epitope by Predict 7 and the latter forming part of epitope 5. The first prediction (Fig. 5.7, panel B1) showed three contacts within epitope 3, 11 in epitope 5 as well as five contacts in one of the polypeptide chains, 87-95. The second EpiPred prediction (Fig. 5.7, panel B2) showed four contacts within epitope 4 (282-299).

The ZDock protocol yields coloured prediction clusters at the most probably positions where the scFv molecule will interact with the antigen. The higher ranked poses are shown in red and change to blue as interaction probability decreases. There were more clusters evident around epitopes 4 and 5 (Fig. 5.7, panels C2 and C3) than for the remaining five epitopes. Since epitope 3 was predicted by EpiPred, these results will be included for completeness (panel C1). The results reported by Huson (2006) were used to select the most probable binding epitope on *Tco*OPB. The affinity purified anti-peptide antibodies raised against epitopes 3, 4 and 5 were all able to detect recombinant *Tco*OPB and *Tvi*OPB. Epitope 5 was poorly immunogenic as low titres of antibodies were produced throughout the immunisation period. The anti-peptide antibodies against both epitopes 3 and 4 were highly immunogenic and high antibody titres were produced. Taken together with the results from the EpiPred (panel B) and ZDock predictions (panel C), scFv interaction at epitope 4 is the most probable scenario.

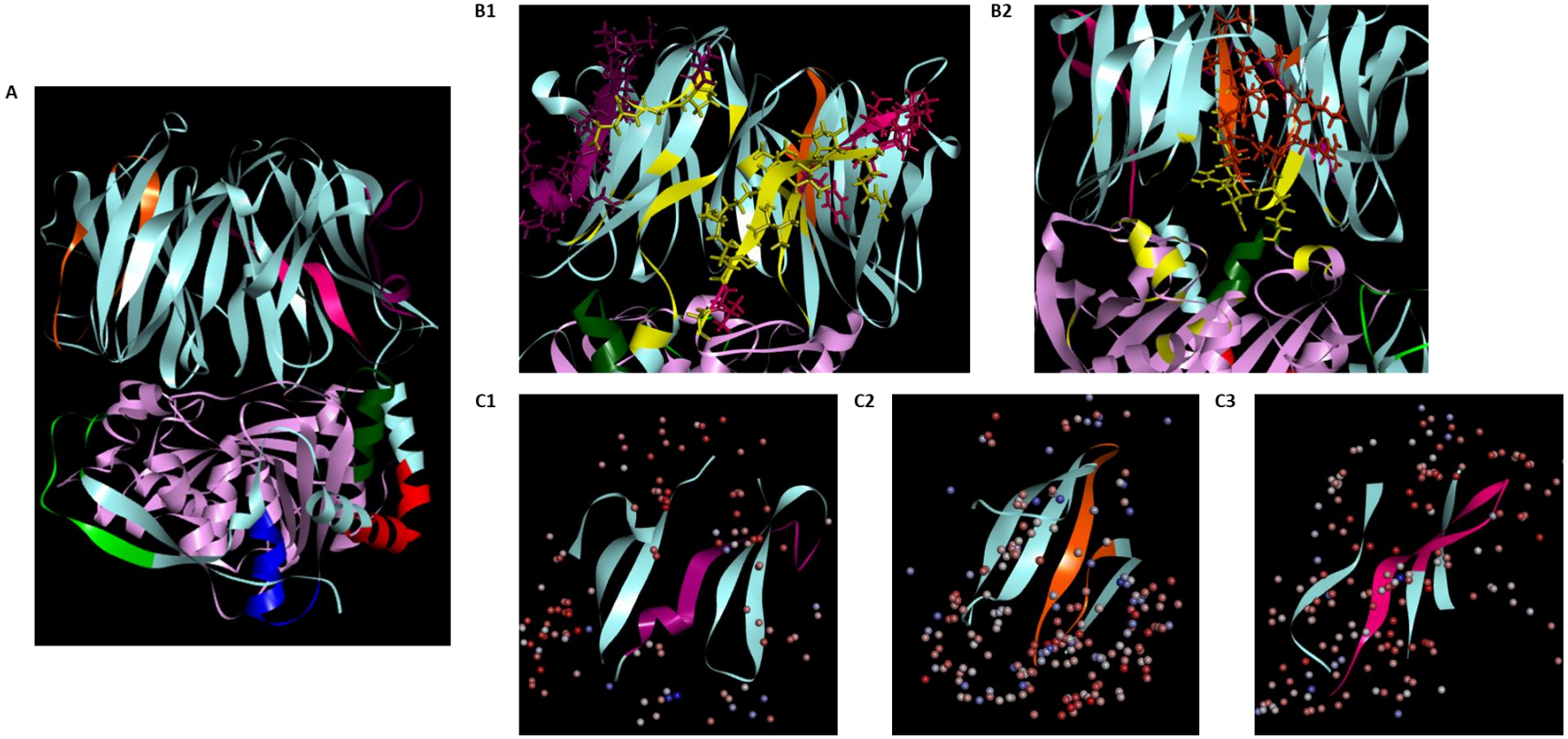


Figure 5.7: Predicted immunogenic and structural epitopes together with the possible scFv interactions with *Tco*OPB. (A) The predicted immunogenic epitopes as predicted by Predict 7 (Cármenes *et al.*, 1989) and reported previously by Huson (2006). Epitopes are coloured in light green (epitope 1), red (2), purple (3), orange (4), pink (5), blue (6) and dark green (7). **(B1 and 2)** The two structural epitope predictions (highlighted in yellow) after entering the E5 scFv V_H and V_L protein sequences together with the 3D structure of *Tco*OPB into the EpiPred predictor (Krawczyk *et al.*, 2014). The amino acid residues of epitopes 3, 4 and 5 are represented in a stick configuration. The ZDock cluster results of the antigen-antibody of E5 scFv and *Tco*OPB around **(C1**, purple**)** epitope 3, **(C2**, orange**)** epitope 4 and **(C3**, pink**)** epitope 5.

5.3.3.4 Docking of scFv onto *TcoOPB*

The modelled structures of *Tco*OPB (Fig. 5.8, panel A), E5 scFv (panel B) were used in the ZDock antibody-antigen interaction protocol using the Discovery Studio software (Dassault Systèmes BIOVIA). The resulting position in which the E5 scFv has been docked onto *Tco*OPB is shown in panel C. As shown in panel D, eight of the 12 V_H CDR3 residues interacted with those of epitope 4. In addition, four residues from each V_H CDR1 and CDR2 and V_L CDR2 interacted with residues from both the catalytic and propeller domain.

5.3.4 Expression of scFv clones

The scFv clones E4, E5 and C2 were all expressed in TG1 *E. coli* cells, which should result in an myc-tagged scFv::pIII fusion protein which is approximately 74 kDa in size (Appendix D1).

A protein band at approximately 35 kDa was identified in the osmotic shock (OS) fraction, more intensely than in the periplasmic (P) fraction, using the affinity purified chicken anti-myc IgY antibody (Fig. 5.9). This protein was present in both the periplasm and osmotic shock fractions for clones E5 and C2, yet only in the osmotic shock fraction for clone E4. This protein may correspond to the myc-tagged scFv without the pIII coat protein, which has an expected size of 31 kDa (Appendix D1). Detection of the 74 kDa pIII fusion protein using the anti-M13 (phage coat protein) antibody was unsuccessful (not shown).

There were no myc-tagged proteins evident in the TG1 *E. coli* cell control. A faint protein band at approximately 21.5 kDa in the osmotic shock fraction was visible for C2 which is smaller than the expected size of myc-tagged scFv without the pIII coat protein.

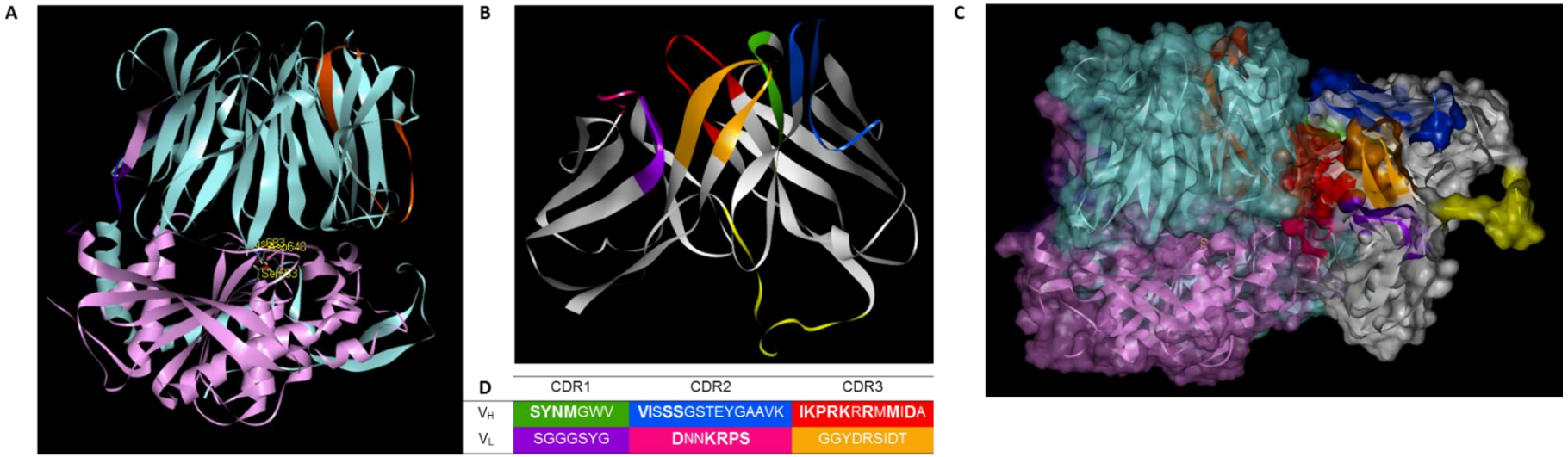


Figure 5.8: Docked scFv onto epitope 4 of *Tco*OPB. (A) Structural features of *Tco*OPB include the catalytic domain (light purple), the propeller domain (light blue), the hinge region of two polypeptide chains (dark blue and dark purple), epitope 4 (orange) and the catalytic triad Ser563, Asp648 and His683 within the catalytic domain. **(B)** The structural features of E5 scFv include the V_H CDR1, -2 and -3 (green, blue, red), the V_L CDR1, -2 and -3 (purple, pink, light orange) and the linker region (bright yellow). **(C)** Solvent accessible surface of docked E5 scFv onto *Tco*OPB using ZDock protocol in the Discovery Studio software (Dassault Systèmes BIOVIA). **(D)** Residues of the E5 scFv CDRs which interact with *Tco*OPB are highlighted in bold.

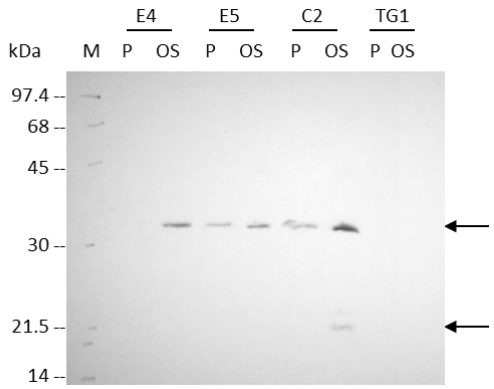


Figure 5.9: Expression of anti-*Tco*OPB and anti-*Tvi*OPB scFv antibodies. Samples of the periplasm (P) and osmotic shock (OS) fractions from the expression of scFv antibodies from *Tco*OPB pan 3 clones E4 and E5 and *Tvi*OPB pan 4 clone C2, all in TG1 *E. coli* cells, were electrophoresed on a 12.5% reducing SDS-PAGE. After transfer onto nitrocellulose, and blocking with 5% (w/v) milk-TBS, the blot was incubated with affinity purified anti-myc IgY [0.5 µg/ml in 0.5% (w/v) BSA-PBS]. Rabbit anti-chicken IgY HRPO conjugate [1:5 000 in 0.5% (w/v) BSA-PBS] and 4-chloro-1-naphthol·H_2O_2 were used as the detection system.

5.3.5 Affinity purification of scFv

Anti-*Tco*OPB scFv (E5 scFv) and anti-*Tvi*OPB scFv (C2 scFv), were purified using the Aminolink® resin, to which purified *Tco*OPB and *Tvi*OPB had been previously coupled. Application of the osmotic shock fraction onto the resin resulted in the purification of a protein, approximately 70 kDa in size, as can be seen by the Coomassie stained gel (Fig. 5.10, panel A). This protein band was also detected by the affinity purified chicken anti-myc IgY antibody (panel B) and corresponds favourably to the predicted molecular weight of the myc-tagged scFv::pIII fusion. Detection of the scFv::pIII fusion protein using the anti-M13 antibody was unsuccessful (not shown).

The fact that the anti-*Tco*OPB E5 scFv and anti-*Tvi*OPB C2 scFv, from the osmotic shock fraction, were able to bind to their respective OPB antigens coupled Aminolink® resins, is direct evidence that the expressed scFvs are specific for *Tco*OPB and *Tvi*OPB. The eluted fractions from the purification of E5 scFv were pooled and used in subsequent experiments.

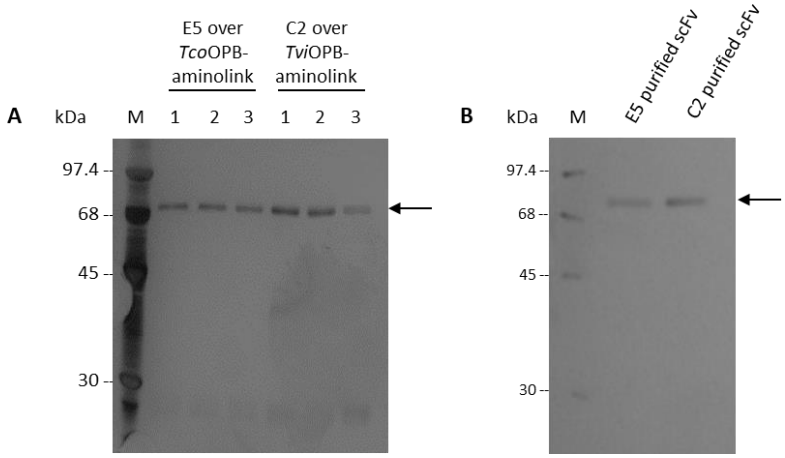


Figure 5.10: Affinity purified scFv using *Tco*OPB-aminolink and *Tvi*OPB-aminolink resins. Samples of the eluted fractions from the purification of E5 anti-*Tco*OPB and C2 anti-*Tvi*OPB scFv from the osmotic shock fraction, using their respective aminolink resins, were electrophoresed on two 12.5% reducing SDS-PAGE gels with one **(A)** stained with Coomassie Blue R-250 and the other transferred onto nitrocellulose, blocked with 5% (w/v) milk-TBS and incubated with **(B)** affinity purified chicken anti-myc IgY [0.5 µg/ml in 0.5% (w/v) BSA-PBS]. Rabbit anti-chicken IgY HRPO conjugate [1:5 000 in 0.5% (w/v) BSA-PBS] and 4-chloro-1-naphthol·H_2O_2 were used as the detection system.

5.3.6 Antigen detection

Since it had been demonstrated that the E5 scFv was able to bind to recombinant *Tco*OPB during purification, the ability of the E5 scFv to detect native *Tco*OPB in the subcellular fractions of *T. congolense* parasites (Section 3.2.5) was tested and compared to the detection by antibodies raised in rabbits and chickens against the recombinant *Tco*OPB and raised in chickens against peptide 4 (Fig. 5.11).

Recombinant *Tco*OPB was detected by each of the four anti-*Tco*OPB antibodies. Native *Tco*OPB is located in the cytosol (Burleigh *et al.*, 1997; Morty *et al.*, 2005a; Rea and Fülöp, 2006), and as such, was detected in both the total parasite proteins and cytosolic fraction, but not in the mitochondrial fraction by each of the four anti-*Tco*OPB antibodies.

Detection of the native *Tco*OPB which was secreted into the culture medium was achieved by three out of the four anti-*Tco*OPB antibodies, each with varying degrees of success. Both the rabbit and chicken anti-*Tco*OPB antibodies detected native *Tco*OPB with a good intensity. However, E5 scFv detection of native *Tco*OPB was very faint, requiring a fair amount of contrast to see the spots. No detection was achieved using the affinity purified chicken anti-peptide 4 IgY antibody.

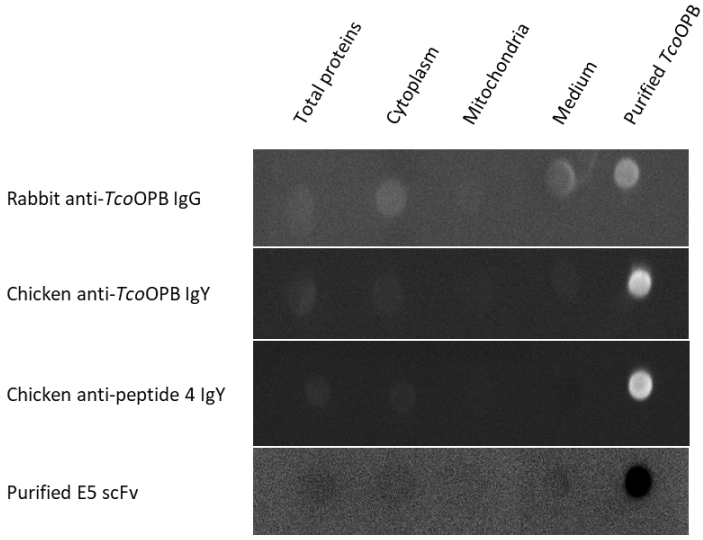


Figure 5.11: Detection of native and recombinant *Tco*OPB. Samples of each *T. congolense* subcellular fraction, together with the precipitated culture medium and recombinant *Tco*OPB, were dotted onto nitrocellulose. After blocking with 5% (w/v) non-fat milk-TBS, the nitrocellulose strips were probed with rabbit anti-*Tco*OPB IgG (10 µg/ml), chicken anti-*Tco*OPB IgY (10 µg/ml), affinity purified chicken anti-peptide 4 IgY (5 µg/ml) and purified scFv (1:25), which were all diluted in 0.5% (w/v) BSA-PBS. Before the addition of the detection antibody, the nitrocellulose strip which had been incubated with the purified E5 scFv, was incubated with affinity purified chicken anti-myc IgY [0.5 µg/ml in 0.5% (w/v) BSA-PBS]. The rabbit anti-chicken IgG HRPO conjugate and the goat anti-rabbit IgG HRPO conjugate, together with the ECL western blotting substrate, were used as the detection system. The panel detected using purified E5 scFv was subject to extensive contrast to better visualise the ECL signal.

As it has been demonstrated that E5 scFv was able to detect native *Tco*OPB, ELISA optimisations were carried out to minimise the amount of scFv used and to increase the detection of *Tco*OPB. Comparison of the ability of the crude E5 scFv, periplasmic and osmotic shock fractions together with the purified scFv, to detect recombinant *Tco*OPB, indicated that the purified E5 scFv was superior (Fig. 5.12, panel A). A 1:100 dilution of the purified E5 scFv yielded in absorbance values of over 1 at *Tco*OPB concentrations as low as 5 µg/ml. At 1 µg/ml, an absorbance of 0.4 was attained.

Absorbance values between 0.210 and 0.356 were obtained when using the periplasmic fraction, even at 10 µg/ml of coated recombinant *Tco*OPB. However, a 1:10 dilution of the osmotic shock fraction produced absorbance values of 0.663 and 0.535 at 10 and 5 µg/ml coated recombinant *Tco*OPB, respectively (Fig. 5.12, panel A). Optimisation of the dilution of the osmotic shock fraction was characterised by absorbance values between 0.222 and 0.263 for the controls in the absence of coated *Tco*OPB, with a 1:2.5 dilution of the osmotic shock fluid having the highest absorbance values at each of the coated recombinant *Tco*OPB concentrations (Fig. 5.12, panel B).

Regrettably, the higher the coated concentration of recombinant *Tco*OPB, the higher the absorbance values of the 'no linking antibody control' (affinity purified chicken anti-myc IgY) and the 'no detection antibody control' (rabbit anti-chicken IgY HRPO conjugate). Despite this, the absorbance values of the osmotic shock fractions were still higher than those of the controls. Several attempts to lower the absorbance values of the controls, through the use of various blocking buffers, and even the use of rabbit anti-myc IgG, were unsuccessful.

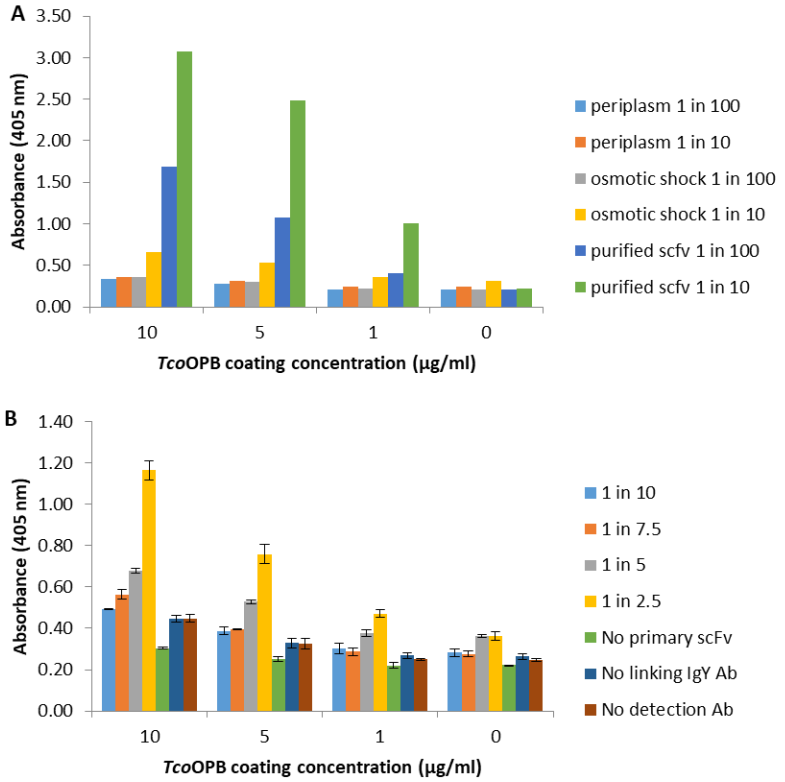


Figure 5.12: Checkerboard ELISA of crude and purified scFv against recombinant *Tco*OPB. ELISA plates were coated with *Tco*OPB (10, 5 and 1 µg/ml in PBS, pH 7.2), blocked with 0.5% (w/v) BSA-PBS and incubated with **(A)** crude scFv, from the periplasm and osmotic shock fractions, and purified scFv (1:10 and 1:100 dilutions) and **(B)** crude scFv from the osmotic shock fraction (1:10, 1:7.5, 1:5 and 1:2.5 dilutions). The bridging antibody used was affinity purified chicken anti-myc IgY (0.5 µg/ml). The rabbit anti-chicken IgY HRPO conjugate (1:5 000) and ABTS·H_2O_2 were used as the detection system. The absorbance readings at 405 nm represent the average of **(A)** duplicate and **(B)** triplicate experiments after 45 min development.

The purified E5 scFv was used to detect native *Tco*OPB in the sera of experimentally *T. congolense* infected cattle using an indirect antigen detection ELISA format. If the rabbit anti-*Tco*OPB antibody was used as the capture antibody, a linking antibody would be required between the scFv and the HRPO conjugate antibody, in the form of the affinity purified chicken anti-myc IgY antibody (Fig. 5.13). This would require a

rabbit anti-chicken HRPO conjugate which may cross-react with the capture antibody. As such, the scFv was used as the capture antibody.

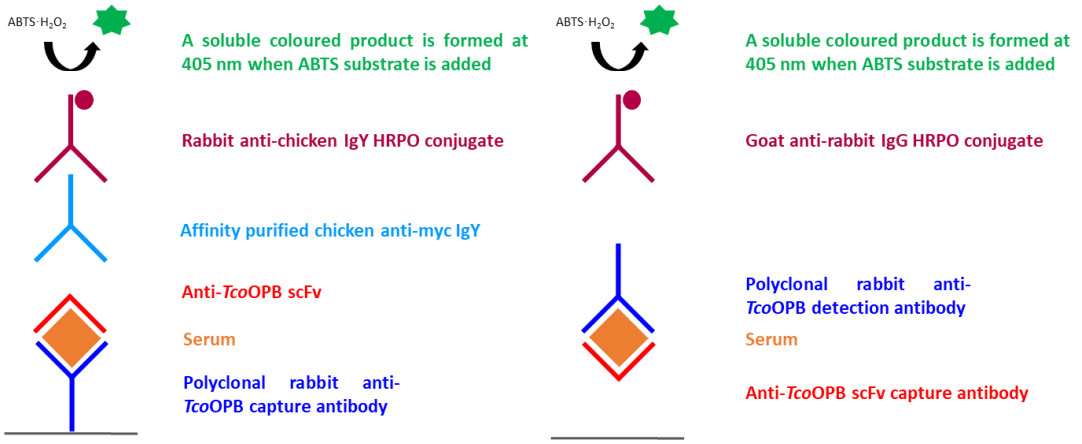


Figure 5.13: Comparison of the choice of capture and detection antibody in the indirect antigen detection ELISA.

Sera from three individual cows, from three separate studies, were selected, based on the reactivity of the sera as tested in previous studies against other possible diagnostic antigens in an antibody detection ELISA format (Eyssen, 2013). A series of serum samples taken from each of these cows, at different times during the infection process, were selected. The values in brackets denote the days prior to and after infection, the T- values denote the days after treatment (Fig. 5.14). Parasitaemia, which was only measured for certain samples by the tecnicians at ClinVet (Bloemfontein), is denoted by black dots. Where parasitaemia was not measured, no value was assigned. Each of the three cows was infected with a different *T. congolense* isolate and treatment was administered after different time periods post infection.

Since OPB is released from dead and dying parasites (Morty *et al.*, 2001), serum samples taken prior to infection should be characterised by low absorbance values in the antigen detection ELISA. During infection, OPB levels should fluctuate together with the characteristic waves of parasitaemia. After successful treatment with trypanocides, dead parasites would release OPB, and result in increased absorbance values in the antigen detection ELISA.

Serum prior to infection had higher absorbance values than that of the recombinant *Tco*OPB and the no serum control, which could suggest that there are factors in the serum samples which may be reacting with the capture and detection antibodies.

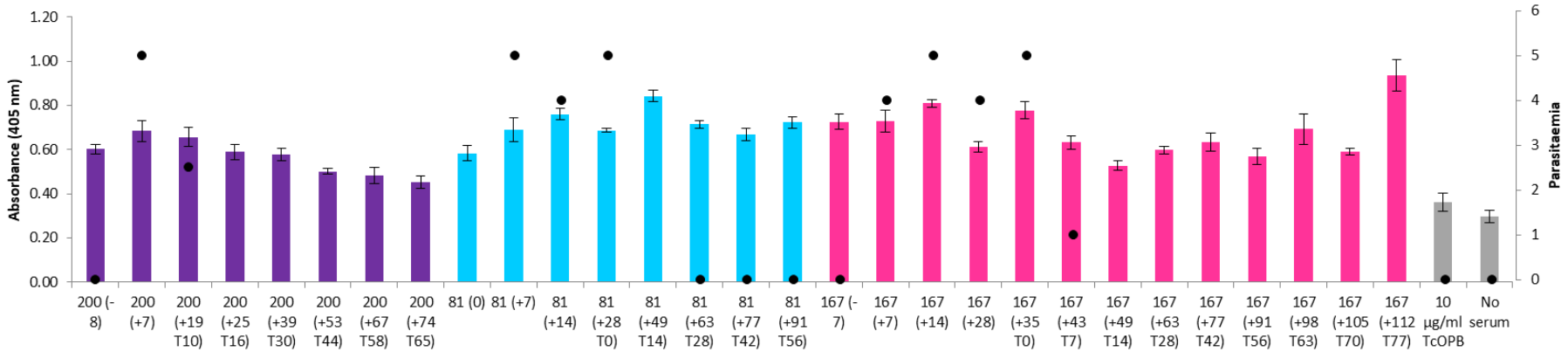


Figure 5.14: Indirect antigen capture ELISA of *Tco*OPB in sera of infected cattle. ELISA plates were coated with the purified scFv (1:100 in PBS, pH 7.2) as the capture antibody, blocked with 0.5% (w/v) BSA-PBS, incubated with sera from infected and non-infected cattle (1:100 dilution), and detected with rabbit anti-*Tco*OPB IgG (1 µg/ml). Goat anti-rabbit IgG HRPO conjugate (1:2 500) and ABTS·H_2O_2 were used as the detection system. The absorbance readings at 405 nm represent the average of triplicate experiments after 30 min development. Parasitaemia is denoted by black dots.

For cow 200, an increase in parasitaemia and absorbance values was evident until day 7 post infection. A steady decrease in absorbance values was observed after treatment, which could indicate that treatment was not effective at killing the parasites, since more OPB would be present in the serum, released by dead parasites, resulting in increased absorbance values. Unfortunately, the parasitaemia was not quantified, but was observed, showing a sustained infection until day 90, 72 days after treatment, before the animal had to be euthanised. This could indicate that due to the ineffective treatment the parasites were not killed, evidenced by the positive confirmation of parasitaemia.

For cow 81, parasitaemia was observed and quantified until day 35, thereafter, it was not evaluated. The ELISA absorbance values increased until day 14, and decreased slightly at day 35. The sharp increase in absorbance values was evident from day 28 to day 48, after 14 days of treatment, which may indicate effective parasite killing where OPB was released into the serum. The next three tested samples had similar absorbance values to that seen at day 35, which could indicate that the treatment was no longer effective.

Similar absorbance values were evident for the sample taken prior to infection and 7 days after, despite an increase in parasitaemia in cow 167. From day 14 to 47, absorbance values fluctuated, together with those of parasitaemia values. After day 49, 14 days post treatment, the absorbance values fluctuate again, peaking at day 112, 77 days post treatment. The overall fluctuation of absorbance values may be linked to that of the fluctuating parasitaemia values, indicating that treatment was not successful as parasites were still evident in the serum and varying levels of OPB was detected throughout the post treatment period.

5.4 Discussion

The serine peptidase of trypanosomal parasites, oligopeptidase B (OPB), was detected in the serum of *T. b. rhodesiense* infected individuals (Eyford *et al.*, 2013). Furthermore, analysis of the secretome of *T. b. gambiense* (Geiger *et al.*, 2010), *T. congolense* (Eyford *et al.*, 2011), *T. cruzi* (Bayer-Santos *et al.*, 2012) and *L. donovani* (Silverman *et al.*, 2008) indicated that OPB from these species are released by the dying parasites. It may therefore be inferred that OPB is an attractive diagnostic antigen. Another antigen shown to be secreted into the host bloodstream is *Tbb*MCA4 (Proto *et al.*, 2011) confirming the detection of *Tbg*MCA4, with which it shares a 100% sequence similarity, in the *T. b. gambiense* secretome (Geiger *et al.*, 2010). The MCAs produced as described in Chapters 2 and 3 showed extensive

autoprocessing with only the double catalytic dyad mutants (*Tco*MCA5^{H147AC202G}) showing minimal processing. However, an insufficient amount of protein could be purified to use to pan the *Nkuku*® phagemid libraries to isolate anti-MCA scFv antibodies. Here the use of phage display to produce recombinant antibodies against OPB as a model antigen for diagnostic application is reported.

Antibodies produced with phage display have been used in a number of applications such as immunoblotting (Renart *et al.*, 1979; Rakabe, 2008), ELISA (Engvall and Perlmann, 1971), immunofluorescence (Coons *et al.*, 1941), flow cytometry (Bonner *et al.*, 1972), affinity chromatography (Porter and Press, 1962) and microarrays (Silzel *et al.*, 1998; Stoevesandt *et al.*, 2009). Phage display has been able to successfully identify antibody fragments using the immunoglobulin genes from camels in a V_HH format (Ghahroudi *et al.*, 1997), from cattle in a Fab format (O'Brien *et al.*, 1999), and from chickens (Davies *et al.*, 1995), rabbits (Ridder *et al.*, 1995) and sheep in an scFv format (Li *et al.*, 2000). Since antibodies are generated from the immunoglobin genes of specific animal hosts, the resulting recombinant antibodies are suitable for veterinary applications. The *Nkuku*® phage display library is based on the immunoglobulin genes from B-cells isolated from the bursa of Fabricus of naïve white leghorn hens (van Wyngaardt *et al.*, 2004). The PCR-amplified genes are incorporated into the pHEN1 phagemid plasmid, the *Nkuku*® library phagemid, which results in the expression of a myc-tagged scFv::pIII fusion protein. When using in the *E. coli* TG1 suppressor strain, the scFv is expressed as the 74 kDa fusion protein.

Panning of the *Nkuku*® library was performed to select *Tco*OPB- and *Tvi*OPB-specific scFv antibodies. The panning process was repeated four times, using the *Nkuku*® library phagemid as the starting material. It was demonstrated by the binding of the anti-M13 (phage coat protein) antibody that enrichment occurred at pans three and four for *Tco*OPB and pan four for *Tvi*OPB. The numerous individual clones selected from each of these pans, were analysed using an ELISA for their ability to produce M13 positive phage binding- and soluble myc positive scFv antibodies. None of the absorbance signals were above 1 and the signals between the phage binding- and soluble ELISA were comparable, unlike the doubling of absorbance values in the soluble scFv ELISA from those of the phage binding ELISA as reported by Rakabe (2008).

Selection of clones for scFv expression is commonly based on the results of the sequencing of the plasmid DNA which becomes costly due to the high number of clones that can be tested. In this study, we employed a selection protocol whereby only the clones which were positive for phage binding and soluble scFv, had complete

coding sequences using colony PCR and those with similar sequences using DNA fingerprinting. This allows for a minimum number of plasmids to be sequenced.

Clones E4 and E5 from *Tco*OPB pan three, and C2 from *Tvi*OPB pan four, were positively identified as positive as complete scFvs and were found to encode for the same protein sequence which may indicate a highly conserved epitope in OPB. The identification of the conserved epitope was achieved using molecular docking of the homology modelled E5 scFv onto *Tco*OPB. This approach has revealed the binding sites of phage displayed scFv on the HSP60 of the *Strongyloides venezuelensis* human infective parasite (Levenhagen *et al.*, 2015). The V_H CDR3 of E5 scFv interacted with *Tco*OPB at a previously identified epitope, VRKREKNVRYEVEMHGT (Huson, 2006). The V_H CDR3 is the main source of antibody variation as it plays a key role in antibody binding (Marks *et al*., 1991; Nissim *et al*., 1994; Sheets *et al*., 1998; van Wyngaardt *et al*., 2004). As such, the interaction of this region with the target antigen verifies the application of the *Tco*OPB-specific scFv in diagnostics. In addition to the CDRs, the framework regions between the CDRs interacted with residues from both the catalytic and propeller domains of *Tco*OPB.

The 3D structure of *Tbb*OPB has been solved, using X-ray diffraction, in both a closed and open conformation, between which a domain rotation of 27.8° is evident (Canning *et al.*, 2013). The process of domain opening disrupts the catalytic triad and results in the inactivation of OPB (Canning *et al.*, 2013). Reactivation requires complete domain closure to restore the catalytic triad, which can only be achieved with peptide substrates which are shorter than 30 residues (Canning *et al.*, 2013). The E5 scFv may not dock in the active site, but may hinder the OPB from adopting the open conformation which would allow substrate access. Due to the difficulties in purifying high concentrations of scFv, immuno-inhibition studies of both recombinant and native *Tco*OPB with E5 scFv could not be carried out for validation of the proposed docking model.

The pelB leader sequence within the pHEN1 plasmid directs the scFv to the periplasm (Kipriyanov *et al.*, 1997). Cold osmotic shock (Neu and Heppel, 1965) was used to obtain the periplasmic and osmotic shock fractions, in which 35 kDa myc-tagged scFv was detected using the anti-myc antibody as has been reported by Philibert *et al*. (2007). Since TG1 *E. coli* cells were used, a 70 kDa fusion protein was expected. Due to the peptidolytic degradation of the myc-tagged scFv::pIII fusion protein in the periplasm, only 20 to 30% of the fusion proteins are intact (McCafferty, 1996; Baek *et al.*, 2002; Shi *et al.*, 2007). When the scFv from the osmotic shock fraction was

concentrated by affinity purification using immobilised OPB, sufficient amounts of intact myc-tagged scFv::pIII fusion protein was available for the anti-myc antibody to detect the full length 70 kDa fusion protein.

Various methods are implemented for the purification of scFv, most of which depend on the affinity tag incorporated onto the scFv by the phagemid plasmid. Both nickel- (Das *et al.*, 2005; Finlay *et al.*, 2005; Saerens *et al.*, 2008; da Silva Ribeiro *et al.*, 2013) and cobalt- (Dong *et al.*, 2013) affinity chromatography has been successfully utilised for purification of 6xHis tagged scFv but the pHEN1 phagemid plasmid only encodes for a myc tag.

The myc-affinity resin has been used to purify scFvs panned from the *Nkuku*® library (van Wyngaardt *et al.*, 2004). However, due to the high cost and difficulties in importing the rabbit anti-myc antibody, two myc affinity resins were made inhouse using both chicken and rabbit anti-myc antibodies coupled to Hydrazide® resin. The myc-antibodies were successful in detecting the scFv in a western blot and ELISA format, but not for the immunoaffinity purification of the expressed scFvs. Purified *Tco*OPB and *Tvi*OPB immobilised on AminoLink® was used to purify the OPB-specific scFvs. The successful purification of the 70 kDa myc-tagged scFv::pIII fusion, indicated that the OPB-specific scFv was indeed able to detect OPB.

Nanobodies (V_HH) from a camelid library, raised against the paraflagellar rod protein in *T. evansi*, detecting the protein in *T. b. brucei, T. congolense* and *T. vivax* (Obishakin *et al.*, 2014), and the *T. congolense* aldolase specific scFv, have both been shown have application in research imaging and diagnostics (Odongo *et al.*, 2016). Despite the ability of scFv to detect recombinant *Tco*OPB, the polyclonal anti-*Tco*OPB antibodies produced in rabbits and chickens were superior in the detection of native *Tco*OPB secreted into the culture medium. This may indicate that the scFv has a lower limit of detection of OPB compared to that of the polyclonal antibodies.

Both the purified and crude E5 scFv were able to detect OPB in a checkerboard ELISA but high background was observed. Cross reactivity was suspected using the polyclonal chicken anti-myc antibody to detect myc-tagged scFv, which is derived from chicken immunoglobulin genes. Rabbits were used to generate anti-myc antibodies but these antibodies did not react as well as those from chickens (results not shown). Blocking with various buffers along with the most commonly used 2% Milk-PBS and 3% BSA-PBS as suggested in the *Nkuku*® library instructions did not reduce the background (results not shown).

The application of scFv antibodies has been successful in the diagnosis of bluetongue virus using an inhibition ELISA (Fehrsen *et al.*, 2005) and human strongyloidiasis using a sandwich ELISA were the immune complexes were detected (Levenhagen *et al.*, 2015). An evaluation of the use of *Tco*OPB in an antibody detection ELISA showed that OPB was not suitable in this format (Eyssen, 2013). As such, neither the inhibition ELISA nor the detection of OPB-immune complexes was pursued.

Using a sandwich ELISA, scFv have been used for the detection of African horse sickness virus, AHVS (van Wyngaardt *et al.*, 2013), infectious bursal disease virus, IBDV (Sapats *et al.*, 2003), *Mycobacterium bovis* (Wemmer *et al.*, 2010), *Neospora caninum* (Dong *et al.*, 2013) and *T. congolense* (Odongo *et al.*, 2016). These ELISAs either made use of a homologous or heterologous scFv pair or a combination of a polyclonal antibody and scFv as the capture and detection antibodies. Modification of the scFv in the pair has been reported by Wemmer *et al*. (2010) were the detection scFv was conjugated to gold particles, and by Odongo *et al*. (2016) where the capture scFv was 6xHis tagged and the detection scFv was biotinylated.

The binding of scFv to the antigen is confirmed using either anti-myc or anti-M13 antibodies. The anti-M13 antibody did not detect the expressed nor the purified E5 scFv in a western blot, and the absorbance values from the phage binding ELISA were similar to those of the soluble scFv ELISA. As such, the anti-myc antibody was used in the detection of the bound scFv when used as a capture antibody.

Using scFv as the capture antibody and polyclonal rabbit anti-*Tco*OPB as the detection antibody, the absorbance values of the serum samples taken at different times during infection, fluctuated together with the parasitaemia values. Since OPB is released into the bloodstream upon parasite lysis, at peak parasitaemia, low levels of OPB is expected and vice versa. Similar to OPB, *Tco*Aldolase is secreted by the parasite (Grébaut *et al.*, 2009) into the bloodstream of the mammalian host by a mechanism which is still unknown. *Tco*Aldolase-specific scFv antibodies used in a homologous sandwich ELISA detected aldolase even after the reduction of parasite numbers (Odongo *et al.*, 2016). It was demonstrated that *Tco*Aldolase-specific scFv was able to distinguish between cured and active infections (Odongo *et al.*, 2016). It is, thus, possible that *Tco*OPB-specific scFv can be used as a diagnostic marker after further adjustments have been made to improve the ELISA performance.

This study showed that using phage display, OPB-specific scFv antibodies were produced that were capable of detecting both recombinant and native *Tco*OPB. These antibodies have potential application in parasite imaging in cellular research and for

diagnostic development. The validation of the immuno-inhibition of OPB requires further investigation along with the generation of *Tco*MCA5-specific scFvs.

CHAPTER 6

GENERAL DISCUSSION

Human African trypanosomiasis (AT) is a neglected tropical disease caused by *Trypanosoma* spp. which can infect both humans, causing human AT (HAT), and animals, causing animal AT (AAT). These haemoprotozoa are transmitted by tsetse flies in sub-Saharan Africa. Human infections occur in 24 countries across the west and central regions of Africa (Simarro *et al.*, 2012). Despite the major impact on livestock farming (Bouyer *et al.*, 2014) and agricultural development in Africa (Alsan, 2015), in contrast to HAT, AAT is not considered as a neglected tropical disease (Roger *et al.*, 2017). In addition, human infective parasites can exist in both wild animals and domestic livestock (Anderson *et al.*, 2011). As most trypanosome infections occur in predominantly rural areas where animal diseases have a direct impact on people's livelihoods (source of food and income), AAT indirectly impacts human health and should be considered as a neglected tropical disease (Roger *et al.*, 2017).

The disability-adjusted life-years (DALYS) caused by HAT has decreased by 74.8% since 2005 (GBD *et al.*, 2017) with fewer than 4 000 human infections reported in 2014 (Franco *et al.*, 2017). The reduction in reported cases is a direct result of active case finding, easier access to healthcare facilities, accurate diagnosis, and timeous treatment (Büscher *et al.*, 2017). Despite their ability to screen and identify potential *T. b. gambiense* infections, the CATT and RDTs are not 100% specific (Jamonneau *et al.*, 2015) and are less specific when disease prevalence is low. The prevalence of *T. b. gambiense* infections is well below 0.1% in most HAT endemic areas, and as a direct result, for a single true positive serological test, 99 false positive tests are reported (Büscher *et al.*, 2017).

The current HAT chemotherapies, which have been in use since the 1960s (Field *et al.*, 2017), have adverse side effects and require lengthy administration at health clinics (Garcia-Salcedo *et al.*, 2014). No new AAT chemotherapies have been introduced since 1961 (Giordani *et al.*, 2016) and drug resistance has been reported in 17 African countries (Van den Bossche and Delespaux, 2011). Due to the immune-evasion mechanisms of African trypanosomes, the development of an effective vaccine is unlikely and as a result, new chemotherapies are required (Field *et al.*, 2017). The identification of the infecting trypanosomal species as well as disease staging is crucial for the treatment of HAT infections. As such, positive serodiagnosis is followed by parasitological diagnosis to confirm infection (Wamboga *et al.*, 2017). This is complicated by the waves of parasitaemia and the resistance of *T. b. gambiense* and

T. b. rhodesiense to the trypanosome lytic factor found in human serum (Vanhollebeke and Pays, 2010). Currently, HAT diagnosis is made using the card agglutination test (CATT), but is being replaced with the simpler BIOLINE HAT RDT in Uganda (Wamboga *et al.*, 2017). This test is better-suited to the lack of cold chain and laboratory facilities in endemic regions. The CATT and RDT meet the ASSURED criteria as they are affordable, sensitive, specific, user friendly, rapid and robust, equipment free and are deliverable to the end user (Peeling *et al.*, 2006). As a result, smaller health care facilities in rural areas are able to offer the BIOLINE RDT test which reduces the distance that the patients need to travel for diagnosis (Wamboga *et al.*, 2017). This allows for more people to be screened and cases to be detected and treated quickly.

Parasitological methods are used for field diagnosis of AAT as well as other tropical diseases (OIE, 2013). Serological diagnosis of AAT includes the indirect fluorescent antibody test and ELISA, both of which do not meet the ASSURED criteria, and are used for disease surveillance (OIE, 2013). Recently, an RDT prototype for the detection of *T. congolense, T. vivax* as well as *T. congolense* and *T. vivax* mixed infections was reported using the recombinant cathepsin-B like peptidase and the tandem repeat protein associated with the flagellum (GM6) antigen (Boulangé *et al.*, 2017). The prototype was characterised by a 92% sensitivity and 95% specificity for *T. congolense* and a 98.2% sensitivity and 95% specificity for *T. vivax* infections.

Two relatively new compounds, fexinidazole and benzoxaborole SCYX-7158, which were identified using phenotypic screening, are in various stages of clinical trials as potential oral HAT chemotherapies (Jacobs *et al.*, 2011; Tarral *et al.*, 2014). Given that infected animals act as reservoirs for human infective trypanosomes (Anderson *et al.*, 2011), together with reports of drug resistance in the absence of new AAT drugs (Van den Bossche and Delespaux, 2011), emphasises the urgent need for new effective AAT chemotherapies.

The development of a vaccine is improbable due to the evasion of the immune system by the parasite as well as elimination of B-cell memory following trypanosome infection (La Greca and Magez, 2011). An anti-disease approach entails targeting factors which are essential to the parasite's survival within either the mammalian host or insect vector, to reduce disease symptoms, rather than killing the parasite per se (Antoine-Moussiaux *et al.*, 2009). The Kinetoplastid Target Database (http://rapid.lifesci.dundee.ac.uk/KTD/) lists potential proteins and peptidases which have been identified as targets for the development of novel chemotherapies against *T. brucei* spp., *T. cruzi* and *Leishmania* spp., the causative agents of African trypanosomiasis, Chagas

disease and Leishmaniasis, respectively. The peptidases studied here, i.e. the MCAs and OPB, are listed in this database as possible drug targets.

The MCAs share the catalytic dyad residues and caspase-haemoglobinase fold (Aravind and Koonin, 2002) with the caspases but differ in numerous aspects, the most important being the MCAs' preference for Arg or Lys at P_1 over Asp. In the early years of MCA research, it was thought that the MCAs functioned in a similar manner to that of the caspases in cell death (Madeo *et al.*, 2002). It has since been reported that the MCAs are sufficiently structurally and functionally diverse to be classed separately from the caspases (McLuskey and Mottram, 2015). Metacaspases have been implicated in cell death processes, but in an indirect manner. To date, the MCAs and caspases share two natural protein substrates: tudor staphylococcal nuclease and glyceraldehyde 3-phosphate dehydrogenase. Both *Picea abies* MCA (Sundström *et al.*, 2009), *Pf*MCA2 (Vandana *et al.*, 2018) and caspase-3 (Sundström *et al.*, 2009) could hydrolyse tudor staphylococcal nuclease, whilst YCA1 (Silva *et al.*, 2011) and caspase-1 (Shao *et al.*, 2007) could hydrolyse glyceraldehyde 3-phosphate dehydrogenase. This indicates that although caspases and MCAs show distinct differences in their P_1 residue requirements, similarities exist in their targeted molecular pathways (Sundström *et al.*, 2009; Bozhkov *et al.*, 2010). Since the MCAs are found in all kingdoms except the metazoa (Uren *et al.*, 2000), and are potentially involved in cell death processes, these peptidases are attractive drug targets.

The hydrolysis of host peptide hormones by OPB, implicates this peptidase in the neurological disturbances of the final stage of AT infection (Morty *et al.*, 1999; Morty *et al.*, 2001). The peptidase is released by dead and dying parasites into the host bloodstream where it remains active (Morty *et al.*, 2005a; Morty *et al.*, 2005b; Munday *et al.*, 2011). Despite being listed in the Kinetoplastid Target Database, it has been shown that OPB is not required for parasite infectivity (Moss *et al.*, 2015) and as such, is not a target for the development of novel chemotherapies. The detection of the released OPB antigen in the mammalian host bloodstream is however, a viable option for the differentiation between cured and active infections. Hence OPB, is an attractive diagnostic antigen and its incorporation into a RDT would be ideal.

The best studied MCAs are those from *A. thaliana* and *S. cerevisiae*. The only multicopy MCAs to be enzymatically characterised are those from *T. b. brucei* (Moss *et al.*, 2007; Proto *et al.*, 2011; Machado *et al.*, 2013) and *T. cruzi* (Kosec *et al.*, 2006), and the single copy MCAs from *L. major* (Gonzáles *et al.*, 2007), *L. donovani* (Lee *et al.*, 2007) and *T. cruzi* (Kosec *et al.*, 2006). The single copy *Tbb*MCA5 was shown to be expressed in both animal and insect infective forms (Helms *et al.*, 2006). The

kinetoplastid single copy MCAs are syntenic orthologues suggesting that they are the possible progenitor of the multicopy MCAs (Ambit *et al.*, 2008). Hence the characterisation of the single copy MCAs from the animal infective *T. congolense* and *T. vivax* could contribute to the understanding of the role played by the MCAs in cellular processes.

In the present study, recombinant *Tbb*MCA2, as well as the MCA5s from *T. congolense* and *T. vivax*, expressed in *E. coli*, were shown to undergo autoprocessing, similar to the caspases (Roy *et al.*, 2014). In addition, immunoaffinity purified native *Tco*MCA showed slight autoprocessing in the current study and this has only been reported previously for native *Lmj*MCA (Gonzáles *et al.*, 2007). Autoprocessing has been observed following recombinant expression of the type I MCAs from *A. thaliana* (Watanabe and Lam, 2005), *Lmj*MCA (Gonzáles *et al.*, 2007), as well as *Tbb*MCA2 (Moss *et al.*, 2007). It is thought that MCA autoprocessing allows for the cleavage of the N- and C-terminal domains, releasing the catalytic domain. Mutation of the residues in *Tbb*MCA2 at which autoprocessing occurred, resulted in an unprocessed enzyme which could hydrolyse small peptide substrates but not large substrates, such as azocasein (Gilio *et al.*, 2017). Comparison of the enzymatic activity of the recombinant full length and the catalytic domain of *Lmj*MCA revealed that the catalytic domain was 3.5 times more active than the full length enzyme (Gonzáles *et al.*, 2007). This could explain the lower rate of Z-Gly-Gly-Arg-AMC hydrolysis by native *Tco*MCA5 than by *Tco*MCA^{H147AC202G} reported in the current study.

Similar to that of *Tbb*MCA2 and *Tvi*MCA5, recombinant expression within insoluble inclusion bodies has been reported for both *Tcr*MCA3 and -5 (Kosec *et al.*, 2006) as well as *Tco*MCA5 (Eyssen, 2013). Nickel affinity chromatography has been used for the purification of MCAs from *Trypanosoma* spp. (Kosec *et al.*, 2006; Moss *et al.*, 2007; Proto *et al.*, 2011), *Leishmania* spp. (Gonzáles *et al.*, 2007; Lee *et al.*, 2007; Peña *et al.*, 2017), *A. thaliana* (Vercammen *et al.*, 2004; Watanabe and Lam, 2011), *S. cerevisiae* (Wong *et al.*, 2012) and *Triticum aestivum* (Piszczek *et al.*, 2012) but the elution patterns obtained in these studies were not reported. In the present study, nickel purification of the sarkosyl solubilised MCAs in the present study resulted in more MCA fragments being eluted than what was loaded onto the resin. This phenomenon was not observed in the immunoaffinity purification of soluble recombinant MCA. This led to the hypothesis that MCA5, which had already undergone autoprocessing during expression, over autoprocessed in the presence of nickel ions. This process was termed nickel-induced over autoprocessing in the present study. The elution profile of the nickel purified MCA5 is similar to that of purified *At*MCA4, which had already

undergone autoprocessing during expression, after the addition of calcium, resulting in the detection of more MCA fragments (Watanabe and Lam, 2011). It was shown that neither the recombinant nor the native *Tco*MCA5 required calcium for their peptidolytic activity. Most type I and -II MCAs are calcium regulated, requiring micromolar and millimolar concentrations of calcium, respectively (Madeo, 2002; Adams, 2003; Wong *et al.*, 2012). Recombinant *A. thaliana* MCA-4, -5 and -8 (He *et al.*, 2008) and *Picea abies* MCA (Bozhkov, 2005) have an absolute calcium requirement for enzymatic activity. It has been demonstrated that recombinant *At*MCA9 (Vercammen *et al.*, 2004), *Lmj*MCA (Gonzáles *et al.*, 2007) and *Tco*MCA5 (Chapter 3), all of which autoprocess, do not require calcium for enzymatic activity. Thus, it can be concluded, that different MCAs have different affinities for different divalent cations.

It has been suggested that recombinant MCA may exhibit minor catalytic activity which could initiate a cascade of peptidolytic processes. Thus, when two proenzyme molecules are in close proximity, one slightly active proenzyme could cleave the second molecule (Watanabe and Lam, 2011). When 6xHis tagged recombinant *Tco*MCA5 and *Tvi*MCA5 bind to the nickel resin, each can interact with itself and those MCA molecules in its immediate surroundings, resulting in processing at secondary cleavage sites as well as Arg and Lys residues. This would result in cleavage within the catalytic domain rendering the peptidase enzymatically inactive.

Mutagenesis of both the catalytic dyad residues of the MCA5s in the present study, resulted in little nickel-induced over autoprocessing, with only the MCA5^{H147AC202G} mutant of *T. congolense* found to be enzymatically active. The activity of the double mutant may be explained by the presence of a secondary catalytic Cys, as reported for *At*MCA9 by Belenghi *et al*. (2007) and the type II MCA from *Triticum aestivum* by Piszczek *et al*. (2012), where the catalytic Cys and secondary catalytic Cys are in close proximity. This hypothesis is strengthened by the conservation of this residue in all kinetoplastid MCAs as reported by Szallies *et al*. (2002) and shown in Appendix A3. In the current study it was demonstrated that the secondary catalytic Cys was hydrogen bonded to one of the conserved water molecules in the active site pocket and formed more Pi-alkyl interactions with the fluorogenic peptide substrates compared to that of the catalytic Cys as determined by docking studies. It has been reported that mutations of the catalytic dyad residues of recombinant *Ldn*MCA were still peptidolytically active, albeit 2-fold less than the wild type peptidase, and were not essential to its 'trypsin-like' activity (Lee *et al.*, 2007). It is, thus, hypothesised that the catalytic His and Cys residues are required for autoprocessing, as evidenced by the limited over

autoprocessing in the double mutant shown in the present study, and that the secondary catalytic Cys is responsible for the peptidolytic activity.

The peptidolytic activity of *Tco*MCA^{H147AC202G} was confirmed by the digestion of gelatin in the present study. It was also shown that the immunoaffinity purified native *Tco*MCA5 was active. This is the first report of the digestion of gelatin by either a caspase or a MCA. The two lower molecular weight forms of *Tco*MCA5^{H147AC202G}, which were thought to contain the catalytic domain, digested gelatin, but the full length protein did not. The full length native enzyme was more active than the single lower molecular weight form which confirms the reports that autoprocessing is not an absolute requirement for peptidolytic activity (Moss *et al.*, 2007).

Reported in this study, *Tco*MCA^{H147AC202G} has a P_1 specificity for Arg and Lys, with a preference for Arg over Lys. Similar to *At*MCA4 (Vercammen *et al.*, 2004) and *At*MCA8 (He *et al.*, 2008), *Pa*MCA (Bozhkov *et al.*, 2005), *Tbb*MCA2 (Machado *et al.*, 2013) and *Ldn*MCA (Lee *et al.*, 2007) which are optimally active between pH 7 and 8.5, *Tco*MCA5^{H147AC202G} was active over a broad pH range between 6 and 9. Inhibition of the enzymatic activity of *Tbb*MCA2 (Moss *et al.*, 2007), *Tbb*MCA4^{S219C} (Proto *et al.*, 2011), *Ldn*MCA (Lee *et al.*, 2007) by antipain and leupeptin was reported and corresponds to the inhibition of *Tco*MCA^{H147AC202G} albeit less effectively, 20% compared to between 90 and 100% inhibition reported for each of the above mentioned MCAs. Inhibition by TLCK (70 to 100%) has been reported for the above mentioned MCAs but did not affect the activity of *Tco*MCA^{H147AC202G}. The irreversible inhibitor Z-Phe-Arg-FMK was used as an active site titrant for the recombinant orthocaspase 1 from *Microcystis aeruginosa* PCC 7806 (Klemenčič *et al.*, 2015), and biotin-Phe-Pro-Arg-CMK inhibited the enzymatic activity of recombinant *At*MCA4 (Watanabe and Lam, 2011). Docking studies performed in the present work, showed that these reversible inhibitors had a lower affinity for *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5 than antipain and leupeptin. Experimental validation of the inhibition of *Tco*MCA^{H147AC202G} is required to confirm this.

The single MCA in *S. cerevisiae*, YCA1 (Madeo, 2002), and the type I MCAs from *A. thaliana* (Vercammen *et al.*, 2004), possess a Gln and Asn-rich N-terminal domain. This domain was found to be responsible for the targeting of the MCA to insoluble protein aggregates as well as the stabilisation of the zymogen (Erhardt *et al.*, 2010; Lee *et al.*, 2010). The YCA1 and single copy kinetoplastid MCAs are phylogenetically more related compared to the multicopy MCAs. The single copy MCAs possess the Pro-, Gln-, Tyr-rich C-terminal extension, which is implicated in protein-protein interactions (Kay *et al.*, 2000). This was demonstrated whereby the overexpression of *Tcr*MCA5 without the C-terminal domain resulted in a drop in parasite cell number and

the presence of non-motile rounded cells resembling apoptotic cells (Laverrière *et al.*, 2012). It is, thus, thought that the C-terminal domain of *Tcr*MCA5 may have pro-apoptotic role, whereby it negatively regulates apoptotic-like cell death. Thus, one can hypothesise that the single copy MCA5s from *T. congolense* and *T. vivax* may function in a similar manner.

Using RNA interference (RNAi), overexpression and gene knock-out studies, reported to date, it has shown that the kinetoplastid MCAs are involved in cell cycle and in the regulation of cell death processes. The RNAi of *Tbb*MCA2, -3 and -5 resulted in the growth arrested parasites with multiple nuclei and kinetoplasts, almost four times more than the non-induced control, indicating an essential role in the cell cycle (Helms *et al.*, 2006). The loss of MCA2 and -3 can be compensated by MCA5 and vise versa as parasites were able to adapt to the successive gene knock-out of MCA2, -3 and -5 (Helms *et al.*, 2006). However, the rapid loss of all three MCA genes by RNAi was lethal to the parasite (Helms *et al.*, 2006).

When overexpressed in yeast, *Tbb*MCA4 was shown to inhibit growth, disrupt mitochondrial function and ultimately resulted in cell death (Szallies *et al.*, 2002). In *T. b. brucei,* MCA4 was found only in the bloodstream trypomastigotes, with RNAi resulting in rapid growth arrest, a decrease in the number of motile cells and the inability to execute cytokinesis (Proto *et al.*, 2011). The deletion of *Tbb*MCA4 prolonged mouse survival by 11 days when compared to that of the WT indicating that *Tbb*MCA4 plays a role in parasite virulence (Proto *et al.*, 2011).

Overexpression of *Tcr*MCA3 in epimastigotes, but not where the catalytic Cys residue had been mutated, resulted in a reduction of parasite growth rate as well as cell division arrest at the G1/S phase (Laverrière *et al.*, 2012. As such, cell division arrest was directly linked to the levels of active enzyme (Laverrière *et al.*, 2012). Likewise, overexpression of *Tcr*MCA3, but not where the catalytic Cys residue had been mutated, was shown to protect epimastigotes from death naturally occurring during the stationary phase and to promote cell differentiation to the metacyclic trypomastigote (Laverrière *et al.*, 2012). Thus, strict regulation of *Tcr*MCA3 is critical for the progression of the *T. cruzi* cell cycle. In addition to the pro-apoptotic role of the C-terminal domain of *Tcr*MCA5, the transfection of *T. cruzi* epimastigotes with *Tcr*MCA5 resulted in a decrease in the number of motile parasites when exposed to fresh human serum (Kosec *et al.*, 2006).

The induction of H_2O_2 cell death in amastigotes resulted in a 2.7-fold increase in *Ldn*MCA activity as well as an increase in cells displaying DNA damage as measured

by TUNEL (Lee *et al.*, 2007). As such, it was hypothesised that the *Ldn*MCAs may act as effector molecules in cell death pathways (Lee *et al.*, 2007). A gene knock-out study on *Ldn*MCA1 showed arrested cell division at the S and G2/M phases, incomplete cytokinesis, and reduced cell growth (Raina and Kaur, 2012). Consequently, a role in the regulation of cell cycle and involvement in cell death was suggested for *Ldn*MCA1.

The overexpression of *Lmj*MCA results in deficiencies in kinetoplast segregation, nuclear division and cytokinesis which could be independent of cell death (Ambit *et al.*, 2008). Higher levels of *Lmj*MCA autoprocessing was observed in actively dividing log phase promastigotes than in stationary promastigotes (Ambit *et al.*, 2008). During cell stress, cytoplasmic *Lmj*MCA is translocated to the mitochondria due to a functional mitochondrial localisation sequence in the N-terminal domain (Zalila *et al.*, 2011). In the mitochondrial matrix the full length *Lmj*MCA is extensively processed into fragments containing the catalytic domain (Zalila *et al.*, 2011). The activity of the catalytic domain enhanced the parasite's sensitivity to H_2O_2 by impairing the mitochondrion (Zalila *et al.*, 2011). This suggests that the catalytic activity of *Lmj*MCA is required for cell death (Gonzáles *et al.*, 2007; Zalila *et al.*, 2011; Casanova *et al.*, 2015).

Using phage display, panning of the recombinant MCA from *L. amazonensis* against a peptide library identified a peptide which shares a high sequence identity to the inhibitor of serine peptidase from *L. major* (Peña *et al.*, 2017). The displayed 7-mer peptide expressed on the N-terminus of the minor phage coat protein (pIII) was shown to inhibit the peptidolytic activity of recombinant *Lam*MCA (Peña *et al.*, 2017). The inhibition of *Lam*MCA by the inhibitor of serine peptidase rather than by the inhibitor of cysteine peptidase supports the finding in the present study that *Tco*MCA5^{H147AC202G} activity is inhibited by serine rather than by cysteine inhibitors. Inhibition of *Lam*MCA with the inhibitor of serine peptidase, shown to bind to MCA, decreased trypsin-like activity and cell death, suggesting that *Lam*MCA is required for heat shock-induced cell death (Peña *et al.*, 2017).

Due to their phylogenetic relatedness and the presence of a Pro-, Gln-, Tyr-rich C-terminal domain, it can be hypothesised that the single copy MCA5s from *T. congolense* and *T. vivax* may function in a similar manner to that of the leishmanial MCAs. RNA interference of a cell cycle enzyme, *Tbb*Aurora kinase 1 (AUK1), demonstrated that the enzyme is essential for parasite virulence and was to be considered as a drug target (Jetton *et al.*, 2009). Through the inhibition of *Tbb*AUK1 activity with a small-molecule inhibitor, VX-680, the functions of AUK1 were elucidated

and was found to be essential for cell cycle progression (Li *et al.*, 2009). Thus, due to their involvement in the cell cycle, MCAs are desirable drug targets.

A peptidomimetic library, designed based on the P_1 specificity of *Tbb*MCA2 for Arg and Lys residues, displayed micromolar IC_{50} values, between 2.1 and 50 μM, when tested in a phenotypic assay against *T. b. brucei, T. cruzi, L. infantum* and *Plasmodium falciparum* (Berg *et al.*, 2012). The IC_{50} values associated with the inhibition of recombinant *Tbb*MCA2 varied between 0.6 and 62 μM (Berg *et al.*, 2012). The 3D structure of *Tbb*MCA2 was solved by X-ray diffraction, two months after the publication of the Berg library (McLuskey *et al.*, 2012). In the present study, the validation of the docking of the designed Berg ligands into the active site of the structure of *Tbb*MCA2, solved by X-ray diffraction, will be performed using molecular docking. This will indicate as to whether or not the reported *in vitro* antiparasitic activity (Berg *et al.*, 2012) is due to the inhibition of *Tbb*MCA2 by the Berg ligands. In addition to the Berg ligands, a commercial library was created consisting of all the peptide substrates and inhibitors reported in the characterisation of other MCAs, and docked into the active site of *Tbb*MCA2 as well as the homology modelled *Tco*MCA5 and *Tvi*MCA5. It was demonstrated that the Berg ligands had high affinity for *Tbb*MCA2 as well as the MCA5s, and that the *Tbb*MCA2 IC_{50} values correlated well to the docking score, indicating that the ligands were specific for the enzyme.

The docking results of the current study highlighted the interactions which were essential for ligand binding. The Berg ligands formed fewer interactions with the residues in the active site than those of the commercial library, and thus, only the commercial ligands were used for the generation of a pharmacophore. This will allow for the screening for new compounds or the optimisation of the Berg ligands to produce a compound with a higher affinity for the MCA substrate in future studies. A similar methodology was employed for the identification of novel inhibitors against the cathepsin-L like peptidase of *T. cruzi*, cruzipain, which is an important virulence factor and drug target (Salas-Sarduy *et al.*, 2017). The GlaxoSmithKline HAT and Chagas chemical boxes were assembled based on the results of phenotypic screening against *T. b brucei* and *T. cruzi* (Peña *et al.*, 2015). Two scaffolds were identified after screening these chemical boxes and used to search the ChEMBL, a database containing information for a large number of drug-like bioactive compounds, to identify new possible ligands. The identified ligands were docked into the active site of cruzipain to determine the interactions between ligand and 'receptor' and the modes of binding (Salas-Sarduy *et al.*, 2017).

Structural water molecules mediate ligand binding, in many cases, through the formation of hydrogen bonds (Ross *et al.*, 2012). Three conserved water molecules, lining the bottom of the active site of *Tbb*MCA2, were shown in the present study to interact with the secondary catalytic Cys92, catalytic His158 and S_1 binding residue Asp211. The Cys92, His158 and Asp211 residues were demonstrated to interact with the ligands in both the Berg and commercial libraries. This supports the notion that these conserved water molecules mediate ligand binding. It has been demonstrated that docking performance is improved by the inclusion of water molecules in both predicted and X-ray diffraction solved structures (Yang and Chen, 2004; de Graaf *et al.*, 2005). This was the case in the current study as more ligands were docked into the active site of the MCAs with higher interaction energies when all the water molecules were included than when they were absent.

Parasite antigens released or secreted into the bloodstream of the mammalian host are ideal for the diagnosis of current infections as anti-trypanosomal antibodies persist after cure (Lejon *et al.*, 2010). The *Tbb*MCA4 is a pseudopeptidase as it possess a Ser in place of the catalytic Cys (Fig. 2.1), and is, thus, catalytically inactive (Proto *et al.*, 2011). This peptidase plays a role in the cell cycle and is also a virulence factor which is secreted into the bloodstream of the mammalian host (Proto *et al.*, 2011). The presence of both OPB and MCA4 in the secretome of *T. b. gambiense* (Geiger *et al.*, 2010) makes OPB and *Tbb*MCA4 attractive targets for AT diagnosis through antigen detection. An ELISA detecting circulating *T. congolense* antigens in goats and cattle, using monoclonal antibodies, was shown to be more sensitive than the microheamocrit centrifugation technique (Nantulya and Lindqvist, 1989; Masake and Nantulya, 1991). Attempts to develop an antigen detection test for AT have shown some success in the diagnosis of experimental infections (Liu *et al.*, 1988; Nantulya and Lindqvist, 1989; Kashiwazaki *et al.*, 1994) but fail when testing samples from natural AT infections (Büscher, 2014). This may be due to low levels of circulating antigens or the concealment of these antigens by immune complexes (Büscher, 2014).

Nanobodies, which are derived from the heavy chain-only antibodies of camels and llamas, are comprised of the V_HH domain, which possesses extended surface loops which are able to penetrate the narrow cavities on various pathogens' surfaces (Nuttall *et al.*, 2004; Streltsov *et al.*, 2004). This was evident by the improved penetration of the nanobody between the variable surface glycoproteins on the surface of *T. b. brucei* compared to the Fab antibody fragment and the anti-VSG IgM antibody (Stijlemans *et al.*, 2004). Nanobodies are characterised by nanomolar affinity for their targets (Skottrup *et al.*, 2011) and the ability to detect both free and bound antigens due to

unique epitope recognition which is different from conventional antibodies (De Genst *et al.*, 2006). Antigen binding antibody fragments can therefore be used to potentially improve antigen detection methods.

The single chain variable fragment (scFv) antibodies, which consist of the V_H and V_L domains, have been successfully used in the development of ELISA immunodiagnostic tests against bluetongue virus (Fehrsen *et al.*, 2005; Rakabe *et al.*, 2011), African horse sickness virus (van Wyngaardt *et al.*, 2004; van Wyngaardt *et al.*, 2013), a 16 kDa antigen of *Mycobacterium tuberculosis* (Sixholo *et al.*, 2011), and the 65 kDa HSP of *Mycobacterium bovis* (Wemmer *et al.*, 2010). The detection of anti-OPB antibodies from sera from cattle experimentally infected with *T. congolense*, using chicken anti-OPB IgY antibodies in an inhibition ELISA format was unable to differentiate between infected and non-infected samples (Eyssen, 2013). As such, a smaller antibody fragment specific for OPB was thought have better chances of detecting OPB, which may be part of immune complexes, in an antigen detection ELISA.

Since the MCAs were demonstrated to over autoprocess when purified using nickel chromatography, OPB was used as the model antigen to optimise the phage display protocol in the present study. The *Nkuku*® phage display library, which was designed using chicken immunoglobulin genes (van Wyngaardt *et al.*, 2004), was used to this end. Panning of the library against both *Tco*OPB and *Tvi*OPB, resulted in the isolation of an OPB-specific scFv antibody. Sequencing of the phagemid plasmid indicated that both the anti-*Tco*OPB and anti-*Tvi*OPB scFv had the same DNA sequence suggesting that the OPB-specific scFv bound to a conserved OPB epitope. Due to the genotype:phenotype linkage, the genetic makeup of the scFv antibody is known and allows for the standardisation and accurate comparison of immunoassays (Schirrmann *et al.*, 2011).

The OPB-specific scFv antibody, purified using a *Tco*OPB affinity resin, was capable of detecting both recombinant and native *Tco*OPB in a western blot format in the current study. The OPB-specific scFv in combination with a polyclonal anti-*Tco*OPB antibody in a sandwich ELISA was utilised for the detection of *Tco*OPB in cattle experimentally infected with *T. congolense*. Despite high background values, which may be as a result of antibody cross reactivity, the absorbance values fluctuated together with the waves of parasitaemia. This indicates that upon parasite killing, low parasitaemia should result, causing higher levels of OPB due to parasite lysis. Prior to optimisation of the antigen detection ELISA in the future, the concentration of OPB in sera at each time point during infection needs to be measured to determine the limit of detection of OPB by scFv. This would confirm if OPB is a bona fida diagnostic

candidate and its application in an antigen detection ELISA using OPB-specific scFv. It has been reported that through the use of a nanobody library from an alpaca vaccinated with the soluble proteome of *T. congolense*, panning against the soluble proteome from four different *T. congolense* strains, an aldolase specific Nb474 was isolated (Odongo *et al.*, 2016). Incorporation of 6xHis tagged and biotin labelled Nb474 in an antigen detection ELISA showed a 87% sensitivity and 94% specificity and could successfully differentiate between active and cured experimental infections (Odongo *et al.*, 2016).

The 3D structure of *Lmj*OPB, solved by X-ray diffraction, for both open and closed structures identified a hinge motion which links the α/β hydrolase catalytic domain from the β-propeller domain (Canning *et al.*, 2013). Domain opening removes Glu172, found in the propeller domain, from the active site, freeing Arg650 which is critical for substrate binding (Canning *et al.*, 2013). The molecular docking of the scFv antibody onto *Tco*OPB reported in the current study, predicted that the scFv bound to a previously identified epitope, residues 282 to 299 (VRKREKNVRYEVEMHGT) as well as residues in both the catalytic and propeller domains. If this is indeed the case, scFv binding would prevent the opening of *Tco*OPB to allow access of the substrate into the catalytic site. Antibody fragments have been reported to inhibit peptidase activity, with inhibition of cancer-associated membrane type serine peptidase (MT-SP) by an scFv (Sun *et al.*, 2003). Panning of a naïve, synthetic, human antibody library against the cancer-associated MT-SP as well as washing with ecotin, a bacterial homologue of the inhibitor of serine peptidase (McGrath *et al.*, 1995), was employed to select for more potent peptidase inhibitors. Several scFv peptide antibodies were isolated of which the most potent had an inhibition kinetic constant of 50 pM against human MT-SP. The immunoinhibitory effect of OPB-specific scFv on the activity of *Tco*OPB requires further investigation.

Since the phage display technology has been shown to produce antigen specific antibodies, it can be applied to the production of MCA-specific antibodies for use either in diagnostics (*Tbb*MCA4), or cellular imaging in the laboratory for the detection of MCA5 and the elucidation of their role in cellular processes using RNAi, overexpression and gene knock-out studies.

In conclusion, the MCA5s from *T. congolense* and *T. vivax* were found to autoprocess and over autoprocess when exposed to nickel ions in the affinity chromatography resin. Mutagenesis studies indicated that the catalytic dyad may be responsible for autoprocessing and that the secondary, putative, catalytic Cys may be responsible for enzymatic activity. Native *Tco*MCA5 was isolated using polyclonal anti-MCA IgY

antibodies and was enzymatically active. Here, we report the first digestion of gelatin by a MCA. The Berg ligands were demonstrated to be inhibitors of *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5 with suggested improvements in binding affinity provided by molecular docking of commercial substrates and inhibitors. Using OPB as a model antigen for the optimisation of the phage display technology, an OPB-specific scFv was isolated which could detect both recombinant and native OPB and has application towards the development of an antigen detection ELISA.

REFERENCES

Abdel-Hamid, M. K., McCluskey, A. (2014). In silico docking, molecular dynamics and binding energy insights into the bolinaquinone-clathrin terminal domain binding site. *Molecules*. **19**, 6609-6622.

Acosta-Maspons, A., Sepúlveda-García, E., Sánchez-Baldoquín, L., Marrero-Gutiérrez , J., Pons, T., Rocha-Sosa, M., González, L. (2014). Two aspartate residues at the putative p10 subunit of a type II metacaspase from *Nicotiana tabacum* L. may contribute to the substrate-binding pocket. *Planta*. **239**, 147-160.

Acup, C., Bardosh, K. L., Picozzi, K., Waiswa, C., Welburn, S. C. (2017). Factors influencing passive surveillance for *T. b. rhodesiense* human African trypanosomiasis in Uganda. *Acta Tropica*. **165**, 230-239.

Adams, J. M. (2003). Ways of dying: multiple pathways to apoptosis. *Genes & development*. **17**, 2481-2495.

Ahmed, H. A., Picozzi, K., Welburn, S. C., MacLeod, E. T. (2013). A comparative evaluation of PCR- based methods for species-specific determination of African animal trypanosomes in Ugandan cattle. *Parasites & Vectors*. **6**, 316.

Aksoy, S., Buscher, P., Lehane, M., Solano, P., Van Den Abbeele, J. (2017). Human African trypanosomiasis control: Achievements and challenges. *PLoS Neglected Tropical Diseases*. **11**, e0005454.

Allsopp, R., Hursey, B. (2004). Insecticidal control of tsetse, In: The Trypanosomiases, (Maudlin, I., Holmes, P., Miles, M. (Eds.)): Wallingford, CABI Publishing.

Alsan, M. (2015). The effect of the TseTse fly on African development. *American Economic Review*. **105**, 382-410.

Alvarez, V., Niemirowicz, G. T., Cazzulo, J. J. (2011). The peptidases of *Trypanosoma cruzi*: digestive enzymes, virulence factors, and mediators of autophagy and programmed cell death. *Biochimica et Biophysica Acta*. **1824**, 195-206.

Ambit, A., Fasel, N., Coombs, G. H., Mottram, J. C. (2008). An essential role for the *Leishmania major* metacaspase in cell cycle progression. *Cell Death & Differentiation*. **15**, 113-122.

Ameisen, J. C., Idziorek, T., Billaut-Mulot, O., Loyens, M., Tissier, J. P., Potentier, A., Ouaissi, A. (1995). Apoptosis in a unicellular eukaryote (*Trypanosoma cruzi*): implications for the evolutionary origin and role of programmed cell death in the control of cell proliferation, differentiation and survival. *Cell Death & Differentiation*. **2**, 285-300.

Anderson, N. E., Mubanga, J., Fèvre, E. M., Picozzi, K., Eisler, M., Thomas, R., Welburn, S. C. (2011). Characterisation of the wildlife reservoir community for human and animal trypanosomiasis in the Luangwa Valley, Zambia. *PLoS Neglected Tropical Diseases*. **5**, e1211.

Antoine-Moussiaux, N., Büscher, P., Desmecht, D. (2009). Host-parasite interactions in trypanosomiasis: on the way to an anti-disease strategy. *Infection & Immunity*. **77**, 1276–1284.

Aravind, L., Koonin, V. E. (2002). Classification of the caspase-hemoglobinase fold: detection of new families and implication for the origin of the eukaryotic separins. *Proteins*. **46**, 355-367.

Ashall, F. (1990). Characterisation of an alkaline peptidase of *Trypanosoma cruzi* and other trypanosomatids. *Molecular & Biochemical Parasitology*. **38**, 77-88.

Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., Depledge, D. P., Fischer, S., Gajria, B., Gao, X., Gardner, M. J., Gingle, A., Grant, G., Harb, O. S., Heiges, M., Hertz-Fowler, C., Houston, R., Innamorato, F., Iodice, J., Kissinger, J. C., Kraemer, E., Li, W., Logan, F. J., Miller, J. A., Mitra, S., Myler, P. J., Nayak, V., Pennington, C., Phan, I., Pinney, D. F., Ramasamy, G., Rogers, M. B., Roos, D. S., Ross, C., Sivam, D., Smith, D. F., Srinivasamoorthy, G., Stoeckert, C. J., Jr., Subramanian, S., Thibodeau, R., Tivey, A., Treatman, C., Velarde, G., Wang, H. (2010). TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucleic Acids Research*. **38**, D457-D462.

Baek, H., Suk, K., Kim, Y., Cha, S. (2002). An improved helper phage system for efficient isolation of specific antibody molecules in phage display. *Nucleic Acids Research*. **30**, e18.

Bagarozzi, D. A., Jr., Potempa, J., Travis, J. (1998). Purification and characterization of an arginine-specific peptidase from ragweed (*Ambrosia artemisiifolia*) pollen. *American Journal of Respiratory Cell & Molecular Biology*. **18**, 363-369.

Baral, T. N. (2010). Immunobiology of African trypanosomes: need of alternative interventions. *Journal of Biomedicine & Biotechnology*. **2010**, 389153.

Barrett, M. P., Boykin, D. W., Brun, R., Tidwell, R. R. (2007). Human African trypanosomiasis: pharmacological reengagement with a neglected disease. *British Journal of Clinical Pharmacology*. **152**, 1155-1171.

Barrett, M. P., Vincent, I. M., Burchmore, R. J., Kazibwe, A. J., Matovu, E. (2011). Drug resistance in human African trypanosomiasis. *Future Microbiology*. **6**, 1037-1047.

Bastos, I. M., Motta, F. N., Grellier, P., Santana, J. M. (2013). Parasite prolyl oligopeptidases and the challenge of designing chemotherapeuticals for Chagas disease, Leishmaniasis and African trypanosomiasis. *Current Medicinal Chemistry*. **20**, 3103-3115.

Bayer-Santos, E., Aguilar-Bonavides, C., Rodrigues, S. P., Cordero, E. M., Marques, A. F., Varela-Ramirez, A., Choi, H., Yoshida, N., da Silveira, J. F., Almeida, I. C. (2012). Proteomic analysis of *Trypanosoma cruzi* secretome: characterization of two populations of extracellular vesicles and soluble proteins. *Journal of Proteome Research*. **12**, 883-897.

Belenghi, B., Romero-Puertas, M. C., Vercammen, D., Brackenier, A., Inzé, D., Delledonne, M., Van Breusegem, F. (2007). Metacaspase activity of *Arabidopsis thaliana* is regulated by *S*-nitosylation of a critical cysteine residue. *Journal of Biological Chemistry*. **282**, 1352-1358.

Benkert, P., Kunzli, M., Schwede, T. (2009). QMEAN server for protein model quality estimation. *Nucleic Acids Research*. **37**, W510-W514.

Benkert, P., Biasini, M., Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*. **7**, 343-350.

Berg, M., Van der Veken, P., Joossens, J., Muthisamy, V., Breugelmans, M., Moss, C. X., Rudolf, J., Cos, P., Coombs, G. H., Maes, L., Haemers, A., Mottram, J. C., Augustyns, K. (2012). Design and evaluation of *Trypanosoma brucei* metacaspase inhibitors. *Bioorganic & Medicinal Chemistry Letters*. **20**, 2001-2006.

Berg, S. S., Brown, K. N., Hill, J., R., W. W. (1961). A new prophylactic trypanocidal drug, 2, 7-bis (M-amidino-pheny ldiazoamino)-10-ethyl-9-phenylphenanthri-dinium chloride di-hydrochloride (M&B 4596). *Nature*. **192**, 365-398.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*. **28**, 235-242.

Berriman, M., Ghedin, E., Hertz-Fowler, C. (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science*. **309**, 416-422.

Bertone, P., Kluger, Y., Lan, N., Zheng, D., Christendat, D., Yee, A., Edwards, A. M., Arrowsmith, C. H., Montelione, G. T., Gerstein, M. (2001). SPINE: an integrated tracking database and data mining approach for identifying feasible targets in high-throughput structural proteomics. *Nucleic Acids Research*. **29**, 2884-2898.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., Bordoli, L., Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*. **42**, W252-W258.

Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S. M., Lee, T., Pope, S. H., Riordan, G. S., Whitlow, M. (1988). Single-chain antigen-binding proteins. *Science*. **242**, 423-426.

Bissantz, C., Kuhn, B., Stahl, M. (2010). A medicinal chemist's guide to molecular interactions. *Journal of medicinal chemistry*. **53**, 5061-5084.

Bisser, S., Lumbala, C., Nguertoum, E., Kande, V., Flevaud, L., Vatunga, G., Boelaert, M., Büscher, P., Josenando, T., Bessell, P. R., Biéler, S., Ndung'u, J. M. (2016). Sensitivity and specificity of a prototype rapid diagnostic test for the detection of *Trypanosoma brucei gambiense* infection: A multi-centric prospective study. *PLoS Neglected Tropical Diseases*. **10**, e0004608.

Blaney, J. (2012). A very short history of structure-based design: How did we get here and where do we need to go? *Journal of Computer-Aided Molecular Design*. **26**, 13-14.

Bleicher, K. H., Bohm, H. J., Muller, K., Alanine, A. I. (2003). Hit and lead generation: beyond high-throughput screening. *Nature Reviews Drug Discovery*. **2**, 369-378.

Block, H., Maertens, B., Spriestersbach, A., Brinker, N., Kubicek, J., Fabis, R., Labahn, J., Schäfer, F. (2009). Immobilized-metal affinity chromatography (IMAC): a review. *Methods in Enzymology*. **463**, 439-473.

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

Blum, M. L., Down, J. A., Gurnett, A. M., Carrington, M., Turner, M. J., Wiley, D. C. (1993). A structural motif in the variant surface glycoproteins of *Trypanosoma brucei*. *Nature*. **362**, 6039-6609.

Bonner, W. A., Hulett, H. R., Sweet, R. G., Herzenberg, L. A. (1972). Fluorescence activated cell sorting. *Review of Scientific Instruments*. **43**, 404-409.

Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J., Schwede, T. (2008). Protein structure homology modeling using SWISS-MODEL workspace. *Nature Protocols*. **4**, 1-13.

Böttger, V., Böttger, A. (2009). Epitope mapping using phage display peptide libraries. *Methods in Molecular Biology*. **529**, 181-201.

Boulangé, A., Pillay, D., Chevtzoff, C., Biteau, N., Come de Graca, 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.

Bourn, D., Grant, B. D., Shaw, A., Torr, S. (2005). Cheap and safe tsetse control for livestock production and mixed farming in Africa. *Aspects of Applied Biology*. **75**, 81.

Bouyer, F., Seck, M. T., Dicko, A. H., Sall, B., Lo, M., Vreysen, M. J., Chia, E., Bouyer, J., Wane, A. (2014). Ex-ante benefit-cost analysis of the elimination of a *Glossina palpalis gambiensis* population in the Niayes of Senegal. *PLOS Neglected Tropical Diseases*. **8**, e3112.

Bouyer, J., Bouyer, F., Donadeu, M., Rowan, T., Napier, G. (2013). Community- and farmer-based management of animal African Trypanosomosis in cattle *Trends in Parasitology*. **29**, 519-522.

Bowden, G. A., Paredes, A. M., Georgiou, G. (1991). Structure and morphology of protein inclusion bodies in *E. coli*. *Biotechnology*. **9**, 725-730.

Bozhkov, P. V. (2005). Cysteine protease mcII-Pa executes programmed cell death during plant embryogenesis. *Proceedings of the National Academy of Sciences USA*. **102**, 14463-14468.

Bozhkov, P. V., Filonova, L. H., Suarez, M. F. (2005). Programmed cell death in plant embryogenesis. *Current Topics in Developmental Biology*. **67**, 135-179.

Bozhkov, P. V., Smertenko, A. P., Zhivotovsky, B. (2010). Apasing out metacaspases and caspases: proteases of many trades. *Science Signalling*. **3**, pe48.

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.

Bressi, J. C., Verlinde, C. L. M., Aronov, A. M., Shaw, M. L., Shin, S. S., Nguyen, L. N., Suresh, S., Buckner, F. S., van Voorhis, W. C., Kuntz, I. D., Hol, W. G. J., Gelb, M. H. (2001). Adenosine analogues as selective inhibitors of glyceraldehyde-3-phosphate dehydrogenase of *Trypanosomatidae* via structure-based drug design. *Journal of Medical Chemistry*. **44**, 2080-2093.

Briggs, G. E., Haldane, J. B. S. (1925). A note on the kinetics of enzyme action. *Biochemical Journal*. **19**, 38-339.

Bruchhaus, I., Roeder, T., Rennenberg, A., Heussler, V. T. (2007). Protozoa parasites: programmed cell death as mechanism of parasitism. *Trends in Parasitology*. **23**, 376-383.

Brun, R., Hecker, H., Lun, Z. R. (1998). *Trypanosoma evansi* and *T. equiperdum*: distribution, biology, treatment and phylogenetic relationship (a review). *Veterinary Parasitology*. **79**, 95-107.

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

Brun, R., Don, R., Jacobs, R. T., Wang, M. Z., Barrett, M. P. (2011). Development of novel drugs for human African trypanosomiasis. *Future Microbiology*. **6**, 677-691.

Burleigh, B. A., Andrews, N. W. (1995). A 120-kDa alkaline peptidase from *Trypanosoma cruzi* is involved in the generation of a novel Ca(2+)-signaling factor for mammalian cells. *Journal of Biological Chemistry*. **270**, 5172-5180.

Burleigh, B. A., Caler, E. V., Webster, P., Andrews, N. W. (1997). A cytosolic serine endopeptidase from *Trypanosoma cruzi* is required for the generation of Ca2+ signaling in mammalian cells. *Journal of Cell Biology*. **136**, 609-620.

Burri, C. (2010). Chemotherapy against human African trypanosomiasis: is there a road to success? *Parasitology*. **137**, 1987-1994.

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

Büscher, P. (2014). Diagnosis of African trypanosomiasis, In: Trypansomes and Trypanosomiasis, (Magez, S., Radwanska, M. (Eds.)): Veinna, Springer.

Büscher, P., Mertens, P., Leclipteux, T., Gilleman, Q., Jacquet, D., Mumba-Ngoyi, D., Pyana, P. P., Boelaert, M., Lejon, V. (2014). Sensitivity and specificity of HAT Sero-K-SeT, a rapid diagnostic test for serodiagnosis of sleeping sickness caused by *Trypanosoma brucei gambiense*: a case-control study. *Lancet Global Health*. **2**, e359-363.

Büscher, P., Cecchi, G., Jamonneau, V., Priotto, G. (2017). Human African trypanosomiasis. *Lancet*. **6736**, 31510-315316.

Caler, E. V., de Avalos, S. V., Haynes, P. A., Andrews, N. W., Burleigh, B. A. (1998). Oligopeptidase B-dependent signalling mediates host cell invasion by *Trypanosoma cruzi*. *EMBO Journal*. **17**, 4975-4986.

Canning, P., Rea, D., Morty, R. E., Fülöp, V. (2013). Crystal structure of *Trypanosoma brucei* oligopeptidase B broaden the paradigm of catalytic regulation in prolyl oligopeptidase family enzymes. *PLoS ONE*. **8**, e79349.

Cardinal, M. V., Reithinger, R., Gürtler, R. E. (2006). Use of an immunochromatographic dipstick test for rapid detection of *Trypanosoma cruzi* in sera from animal reservoir hosts. *Journal of Clinical Microbiology*. **44**, 3005-3007.

Cármenes, R. S., Freije, J. P., Molina, M. M., Martín, J. M. (1989). Predict 7, a program for protein structure prediction. *Biochemical & Biophysical Research Communications*. **159**, 687-693.

Casanova, M., Gonzalez, I. J., Sprisser, C., Zalila, H., Dacher, M., Basmaciyan, L., Späth, G. F., Azas, N., Fasel, N. (2015). Implication of different domains of the *Leishmania major* metacaspase in cell death and autophagy. *Cell death & disease*. **6**, e1933.

Castanys-Muñoz, E., Brown, E., Coombs, G. H., Mottram, J. C. (2012). *Leishmania mexicana* metacaspase is a negative regulator of amastigote proliferation in mammalian cells. *Cell death & disease*. **3**, e385.

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

Checchi, F., Filipe, J. A. N., Haydon, D. T., Chandramohan, D., Chappuis, F. (2008). Estimates of the duration of the early and late stage of gambiense sleeping sickness. *BMC Infections Diseases*. **8**, 16.

Chema, D., Eren, D., Yayon, A., Goldblum, A., Zaliani, A. (2004). Identifying the binding mode of a molecular scaffold. *Journal of Computer-Aided Molecular Design*. **18**, 23-40.

Chitanga, S., Marcotty, T., Namangala, B., Van den Bossche, P., Van Den Abbeele, J., Delespaux, V. (2011). High prevalence of drug resistance in animal trypanosomes without a history of drug exposure. *PLoS Neglected Tropical Diseases*. **5**, e1454.

Choi, C. J., Berges, J. A. (2013). New types of metacaspases in phytoplankton reveal diverse origins of cell death proteases. *Cell Death & Differentiation*. **4**, e490.

Chowdhury, P. S., Viner, J. L., Beers, R., Pastan, I. (1998). Isolation of a high-affinity stable single-chain Fv specific for mesothelin from DNA-immunized mice by phage display and construction of a recombinant immunotoxin with anti-tumor activity. *Proceedings of the National Academy of Sciences of the United States of America*. **95**, 669-674.

Christendat, D., Yee, A., Dharamsi, A., Kluger, Y., Savchenko, A., Cort, J. R., Booth, V., Mackereth, C. D., Saridakis, V., Ekiel, I., Kozlov, G., Maxwell, K. L., Wu, N., McIntosh, L. P., Gehring, K., Kennedy, M. A., Davidson, A. R., Pai, E. F., Gerstein, M., Edwards, A. M., Arrowsmith, C. H. (2000). Structural proteomics of an archaeon. *Natural Structural Biology*. **7**, 903-909.

Coetzer, T. H. T., Goldring, J. P., Huson, L. E. J. (2008). Oligopeptidase B: A processing peptidase involved in pathogenesis. *Biochimie*. **90**, 336-344.

Cohen, S. N., Chang, A. C. Y., Hsu, L. (1972). Nonchromosomal antibiotoc resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proceedings of the National Academy of Sciences*. **69**, 2110-2114.

Coll, N. S., Vercammen, D., Smidler, A., Clover, C., Van Breusegem, F., Dangl, J. L., Epple, P. (2010). *Arabidopsis* type I metacaspases control cell death. *Science*. **330**, 1393-1397.

Conroy, P. J., O'Kennedy, R. J., Hearty, S. (2012). Cardiac troponin I: a case study in rational antibody design for human diagnostics. *Protein Engineering, Design & Selection*. **25**, 395-305.

Conroy, P. J., Law, R. H., Gilgunn, S., Hearty, S., Caradoc-Davies, T. T., Lloyd, G., O'Kennedy, R. J., Whisstock, J. C. (2014). Reconciling the structural attributes of avian antibodies. *Journal of Biological Chemistry*. **289**, 15384-15392.

Coons, A. H., Creech, H. J., Jones, R. N. (1941). Immunological properties of an antibody containing a fluorescence group. *Proceedings of the Society for Experimental Biology & Medicine*. **47**, 200-202.

Cordon-Obras, C., Berzosa, P., Ndong-Mabale, N., Bobuakasi, L., Buatiche, J. N., Ndongo-Asumu, P., Benito, A., Cano, J. (2009). *Trypanosoma brucei gambiense* in domestic livestock of Kogo and Mbini foci (Equatorial Guinea). *Tropical Medicine & International Health*. **14**, 535-541.

Courtin, F., Camara, M., Rayaisse, J. B., Kagbadouno, M., Dama, E., Camara, O., Traore, I. S., Rouamba, J., Peylhard, M., Somda, M. B., Leno, M., Lehane, M. J., Torr, S. J., Solano, P., Jamonneau, V., Bucheton, B. (2015). Reducing human-tsetse contact significantly enhances the efficacy of sleeping sickness active screening campaigns: A promising result in the context of elimination. *PLoS Neglected Tropical Diseases*. **9**, e0003727.

Cross, G. A. M., Kim, H., Wickstead, B. (2014). Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Molecular & Biochemical Parasitology*. **195**, 59-73.

Crowe, J. S., Lamont, A. G., Barry, J. D., Vickerman, K. (1984). Cytotoxicity of monoclonal antibodies to *Trypanosoma brucei*. *Transactions of the Royal Society of Tropical Medicine & Hygiene*. **78**, 508-513.

da Silva Ribeiro, V., Araújo, T. G., Gonzaga, H. T., Nascimento, R., Goulart, L. R., Costa-Cruz, J. M. (2013). Development of specific scFv antibodies to detect neurocysticercosis antigens and potential applications in immunodiagnosis. *Immunology Letters*. **156**, 59-67.

Das, D., Allen, T. M., Suresh, M. R. (2005). Comparative evaluation of two purifcation methods of anti-CD19-c-myc-His6-Cys scFv. *Protein Expression & Purification*. **39**, 199-208.

Davies, E. L., Smith, J. S., Birkett, C. R., Manser, J. M., Anderson-Dear, D. V., Young, J. R. (1995). Selection of specific phage-display antibodies using libraries derived from chicken immunoglobulin genes. *Journal of Immunological Methods*. **186**, 125-135.

de Andrade, A. S., Santoro, M. M., de Melo, M. N., Mares-Guia, M. (1998). *Leishmania (Leishmania) amazonensis*: purification and enzymatic characterization of a soluble serine oligopeptidase from promastigotes. *Experimental Parasitology*. **89**, 153-160.

De Genst, E., Silence, K., Ghahroudi, M. A., Decanniere, K., Loris, R., Kinne, J., Wyns, L., Muyldermans, S. (2005). Strong in vivo maturation compensates for structuraly restricted H3 loops in antibody repertoires. *Journal of Biological Chemistry*. **280**, 14114-14121.

De Genst, E., Silence, K., Decanniere, K., Conrath, K., Loris, R., Kinne, J., Muyldermans, S., Wyns, L. (2006). Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proceedings of the National Academy of Sciences of the United States of America*. **103**, 4586-4591.

de Graaf, C., Pospisil, P., Pos, W., Folkers, G., Vermeulen, N. P. E. (2005). Binding mode prediction of cytochrome P450 and thymidine kinase protein-ligand complexes by consideration of water and rescoring in automated docking. *Journal of medicinal chemistry*. **48**, 2308-2318.

De Greef, C., Imberechts, H., Matthyssens, G., Van Meirvenne, N., Hamers, R. (1989). A gene expressed only in serum-resistant variants of *Trypanosoma brucei rhodesiense*. *Molecular & Biochemical Parasitology*. **36**, 169-176.

Deckert, P. M. (2009). Current constructs and targets in clinical development for antibody-based cancer therapy. *Current Drug Targets*. **10**, 158-175.

Delespaux, V., Geerts, S., Brandt, J., Elyn, R., Eisler, M. C. (2002). Monitoring the correct use of isometamidium by farmers and veterinary assistants in Eastern Province of Zambia using isometamidium-ELISA. *Veterinary Parasitology*. **110**, 117-122.

Delespaux, V., de Koning, H. (2007). Drugs and drug resistance in African trypanosomiasis. *Drug Resistance Updates*. **10**, 30-50.

Dempsey, W. A., Mansfield, J. M. (1983). Lymphocyte function in experimental African trypanosomiasias. V. Role of antibody and the mononuclear phagocyte system in variant specific immunity. *Journal of Immunology*. **130**, 405-411.

Deponte, M. (2008). Programmed cell death in protists. *Biochimica et Biophysica Acta*. **1783**, 1396-1405.

Desquesnes, M., McLaughlin, G., Zoungrana, A., Davila, A. M. (2001). Detection and identification of Trypanosoma of African livestock through a single PCR based on internal transcribed spacer 1 of rDNA. *International Journal for Parasitology*. **31**, 610-614.

Diall, O., Cecchi, G., Wanda, G., Argiles-Herrero, R., Vreysen, M. J. B., Cattoli, G., Viljoen, G. J., Mattioli, R., Bouyer, J. (2017). Developing a progressive control pathway for African animal trypanosomosis. *Trends in Parasitology*. **33**, 499-509.

Dong, J., Otsuki, T., Kato, T., Kohsaka, T., Ike, K., Park, E. Y. (2013). Development of two murine antibodies against *Neospora caninum* using phage display technology and application on the detection of *N. caninum PLoS ONE*. **8**, e53264.

Dong, L., Masaki, Y., Takegami, T., Kawanami, T., Itoh, K., Jin, Z., Huang, C., Tong, X., Fukushima, T., Tanaka, M., Sawaki, T., Sakai, T., Sugai, S., Okazaki, T., Hirose, Y.,

Umehara, H. (2007). Cloning and expression of two human recombinant monoclonal Fab fragments specific for EBV viral capsid antigen. *International Immunology*. **19**, 331-336.

Drag, M., Salvesen, G. S. (2010). Emerging principles in protease-based drug discovery. *Nature Reviews Drug Discovery*. **9**, 690-701.

Dunbar, J., Krawczyk, K., Leem, J., Marks, C., Nowak, J., Regep, C., Georges, G., Kelm, S., Popovic, B., Deane, C. M. (2016). SAbPred: a structure-based antibody prediction server. *Nucleic Acids Research*. **44**, W474-W478.

Dussart, P., Petit, L., Labeau, B., Bremand, L., Leduc, A., Moua, D., Matheus, S., Baril, L. (2008). Evaluation of two new commercial tests for the diagnosis of acute dengue virus infection using NS1 antigen detection in human serum. *PLoS Neglected Tropical Diseases*. **2**, e280.

Duxbury, R. E., Sadun, E. H., Wellde, B. T., Anderson, J. S., Muriithi, I. E. (1972). Immunization of cattle with x-irriadiated African trypanosomes. *Transactions of the Royal Society of Tropical Medicine & Hygiene*. **66**, 349-350.

Eisenthal, R., Danson, M. J., Hough, D. W. (2007). Catalytic efficiency and k_{cat}/K_M: a useful comparator? *Trends in Biotechnology*. **25**, 247-249.

Ellis, K. J., Morrison, J. F. (1982). Buffers of constant ionic strength for studying pH-dependent processes. *Methods in Enzymology*. **87**, 405-426.

Emmer, B. T., Souther, C., Toriello, K. M., Olson, C. L., Epting, C. L., Engman, D. M. (2009). Identification of a palmitoyl acyltransferase required for protein sorting to the flagellar membrane. *Journal of Cell Science*. **122**, 867-874.

Engstler, M., Thilo, L., Weise, F., Grunfelder, C. G., Schwarz, H., Boshart, M., Overath, P. (2004). Kinetics of endocytosis and recycling of the GPI-anchored variant surface glycoprotein in *Trypanosoma brucei*. *Journal of Cell Science*. **117**, 1105-1115.

Engvall, E., Perlmann, P. (1971). Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. *Immunochemistry*. **8**, 871-874.

Erhardt, M., Wegrzyn, R. D., Deuerling, E. (2010). Extra N-terminal residues have a profound effect on the aggregation properties of the potential yeast prion protein Mca1. *PLoS ONE*, e9929.

Eyford, B. A., Sakurai, T., Smith, D., Loveless, B., Hertz-Fowler, C., Donelson, J. E., Inoue, N., Pearson, T. W. (2011). Differential protein expression throughout the life cycle of *Trypanosoma congolense*, a major parasite of cattle in Africa. *Molecular & Biochemical Parasitology*. **177**, 116-125.

Eyford, B. A., Ahmad, R., Enyaru, J. C., Carr, S. A., Pearson, T. W. (2013). Identification of trypanosome proteins in plasma from African sleeping sickness patients infected with *T. b. rhodesiense*. *PLoS ONE*. **8**, e71463.

Eyssen, L. E., (2013). Studying trypanosomal peptidase antigen targets for the diagnosis of African trypanosomiasis. University of KwaZulu-Natal, Pietermaritzburg. MSc.

Fehrsen, J., van Wyngaardt, W., Mashau, C., Potgieter, A. C., Chaudhary, V. K., Gupta, A., Jordaan, F. A., Du Plessis, D. H. (2005). Serogroup-reactive and type specific detection of bluetongue virus antibodies using chicken scFvs in inhibition ELISAs. *Journal of Virological Methods*. **129**, 31-39.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*. **39**, 783-791.

Fernandes, L. C., Bastos, I. M., Lauria-Pires, L., Rosa, A. C., Teixeira, A. R., Grellier, P., Schrevel, J., Santana, J. M. (2005). Specific human antibodies do not inhibit *Trypanosoma cruzi* oligopeptidase B and cathepsin B, and immunoglobulin G enhances the activity of trypomastigote-secreted oligopeptidase B. *Microbes & Infection*. **7**, 375-384.

Ferreira, L. G., dos Santos, R. N., Oliva, G., Andricopulo, A. D. (2015). Molecular docking and structure-based drug design strategies. *Molecules*. **20**, 13384-13421.

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

Févre, E. M., Wissmann, B. V., Welburn, S. C., Lutumba, P. (2008). The burden of human African trypanosomiasis. *PLoS Neglected Tropical Diseases*. **2**, e333.

Field, M. C., Carrington, M. (2009). The trypanosome flagellar pocket. *Nature Reviews Microbiology*. **7**, 775-786.

Field, M. C., Lumb, J. H., Adung'a, V. O., Jones, N. G., Engstler, M. (2009). Macromolecular trafficking and immune invasion in African trypanosomes. *International Review of Cell & Molecular Biology*. **278**, 1-67.

Field, M. C., Horn, D., Fairlamb, A. H., Ferguson, M. A. J., Gray, D. W., Read, K. D., De Rycker, M., Torrie, L. S., Wyatt, P. G., Wyllie, S., Gilbert, I. H. (2017). Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need. *Nature Reviews Microbiology*. **15**, 217-231.

Finlay, W. J., deVore, N. C., Dobrovolskaia, E. N., Gam, A., Goodyear, C. S., Slater, J. E. (2005). Exploiting the avian immunoglobulin system to simplify the generation of recombinant antibodies to allergenic proteins. *Clinical & Experimental Allergy*. **35**, 1040-1048.

Fleming, J. R., Sastry, L., Wall, S. J., Sullivan, L., Ferguson, M. A. J. (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**, e0004977.

Foloppe, N., Fisher, L. M., Howes, R., Potter, A., Robertson, A. G., Surgenor, A. E. (2006). Identification of chemically diverse Chk1 inhibitors by receptor-based virtual screening. *Bioorganic & Medicinal Chemistry*. **14**, 4792-4802.

Fraga, J., Fernandez-Calienes, A., Montalvo, A. M., Maes, I., Deborggraeve, S., Buscher, P., Dujardin, J. C., Van der Auwera, G. (2016). Phylogenetic analysis of the *Trypanosoma* genus based on the heat-shock protein 70 gene. *Infection, Genetics & Evolution*. **43**, 165-172.

Franco, J. R., Cecchi, G., Priotto, G., Paone, M., Diarra, A., Grout, L., Mattioli, R., Argaw, D. (2017). Monitoring the elimination of human African trypanosomiasis: Update to 2014. *PLoS Neglected Tropical Diseases*. **11**, e0005585.

Frearson, J. A., Wyatt, P. G., Gilbert, I. H., Fairlamb, A. H. (2007). Target assessment for antiparasitic drug discovery. *Trends in Parasitology*. **23**, 589-595.

Fuentes-Prior, P., Salvesen, G. S. (2004). The protein structures that shape caspase activity, specificity, activation and inhibition. *Biochemical Journal*. **384(Pt 2)**, 201-232.

Fyfe, J., Picozzi, K., Waiswa, C., Bardosh, K. L., Welburn, S. C. (2017). Impact of mass chemotherapy in domestic livestock for control of zoonotic *T. b. rhodesiense* human African trypanosomiasis in Eastern Uganda. *Acta Tropica*. **165**, 216-229.

Gannavaram, S., Debranbant, A. (2012). Programmed cell death in Leshimania: biochemical evidence and role in parasite infectivity. *Frontiers in Cellular & Infection Microbiology*. **2**, 95.

Garcia-Salcedo, J. A., Munday, J. C., Unciti-Broceta, J. D., de Koning, H. P. (2014). Progress towards new treatments for human African trypanosomiasis, In: Trypanosomes & Trypanosomiasis, (Magez, S., Radwanska, M. (Eds.)): Vienna, Springer.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server, In: The Proteomics Protocols Handbook, (Walker, J. M. (Ed.), Humana Press.

GBD, DALYS, HALE, Collaborators. (2017). Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016 *Lancet*. **390**, 1260-1344.

Geerts, S., Holmes, P. H., Diall, O., Eisler, M. C. (2001). African bovine trypanosomiasis: the problem of drug resistance. *Trends in Parasitology*. **17**, 25-28.

Geiger, A., Hirtz, C., Bécue, T., Bellard, E., Centeno, D., Gargani, D., Rossignol, M., Cuny, G., Peltier, J. B. (2010). Exocytosis and protein secretion in *Trypanosoma*. *BMC Microbiology*. **10**, 20.

Geysen, D., Delespaux, V., Geerts, S. (2003). PCR–RFLP using Ssu-rDNA amplification as an easy method for species-specific diagnosis of Trypanosoma species in cattle. *Veterinary Parasitology*. **110**, 171-180.

Ghahroudi, M. A., Desmyter, A., Wyns, L., Hamers, R., Muyldermans, S. (1997). Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Letters*. **414**, 521-526.

Gilbert, I. H. (2013). Drug discovery for neglected diseases: molecular target-based and phenotypic approaches. *Journal of medicinal chemistry*. **56**, 7719-7726.

Gilio, J. M., Marcondes, M. F., Ferrari, D., Juliano, M. A., Juliano, L., Oliveira, V., Machado, M. F. (2017). Processing of metacaspase 2 from *Trypanosoma brucei* (TbMCA2) broadens its substrate specificity. *Biochimica et Biophysica Acta*. **1865**, 388-394.

Gilson, M. K., Liu, T., Baitaluk, M., Nicola, G., Hwang, L., Chong, J. (2016). BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Research*. **44**, D1045-D1053.

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.

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

Goldring, J. P. D., Thobakgale, C., Hiltunen, T., Coetzer, T. H. T. (2005). Raising antibodies in chickens against primaquine, pyrimethamine, dapsone, tetracycline, and doxycycline. *Immunological Investigations*. **34**, 101-114.

Gonzáles, I. J., Desponds, C., Schaff, C., Mottram, J. C., Fasel, N. (2007). *Leishmania major* metacaspase can replace yeast metacaspase in programmed cell death and has arginine-specific cysteine peptidase activity. *International Journal for Parasitology*. **37**, 161-172.

Gonzatti, S. M., González-Baradat, B., Aso, P. M., Reyna-Bello, A. (2014). *Trypanosoma (Duttonella) vivax* and Typanosomosis in Latin America: Secadera/Huequera/Cacho Hueco, In: Trypanosomes and Trypanosomiasis, (Magez, S., Radwanska, M. (Eds.)): Vienna, Springer-Verlag.

Goodarzi, M., da Cunha, E. F. F., Freitas, M. P., Ramalho, T. C. (2010). QSAR and docking studies of novel antileishmanial diaryl sulfides and sulfonamides. *European Journal of Medicinal Chemistry*. **45**, 4879-4889.

Grant, I. F. (2001). Insecticides for tsetse and trypanosomiasis control: is the environmental risk acceptable? *Trends in Parasitology*. **17**, 10-14.

Graslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., Ming, J., dhe-Paganon, S., Park, H. W., Savchenko, A., Yee, A., Edwards, A., Vincentelli, R., Cambillau, C., Kim, R., Kim, S. H., Rao, Z., Shi, Y., Terwilliger, T. C., Kim, C. Y., Hung, L. W., Waldo, G. S., Peleg, Y., Albeck, S., Unger, T., Dym, O., Prilusky, J., Sussman, J. L., Stevens, R. C., Lesley, S. A., Wilson, I. A., Joachimiak, A., Collart, F., Dementieva, I., Donnelly, M. I., Eschenfeldt, W. H., Kim, Y., Stols, L., Wu, R., Zhou, M., Burley, S. K., Emtage, J. S., Sauder, J. M., Thompson, D., Bain, K., Luz, J., Gheyi, T., Zhang, F., Atwell, S., Almo, S. C., Bonanno, J. B., Fiser, A., Swaminathan, S., Studier, F. W., Chance, M. R., Sali, A., Acton, T. B., Xiao, R., Zhao, L., Ma, L. C., Hunt, J. F., Tong, L., Cunningham, K., Inouye, M., Anderson, S., Janjua, H., Shastry, R., Ho, C. K., Wang, D., Wang, H., Jiang, M., Montelione, G. T., Stuart, D. I., Owens, R. J., Daenke, S., Schutz, A., Heinemann, U., Yokoyama, S., Bussow, K., Gunsalus, K. C. (2008). Protein production and purification. *Nature methods*. **5**, 135-146.

Grébaut, P., Chuchana, P., Brizard, J. P., Demettre, E., Seveno, M., Bossard, G., Jouin, P., Vincendeau, P., Bengaly, Z., Boulangé, A., Cuny, G., Holzmuller, P. (2009). Identification of total and differentially expressed excreted-secreted proteins from *Trypanosoma congolense* strains exhibiting different virulence and pathogenicity. *International Journal for Parasitology*. **39**, 1137-1150.

Gupta, S., Kim, C., Yel, L., Gollapudi, S. (2004). A role of fas-associated death domain (FADD) in increased apoptosis in aged humans. *Journal of clinical immunology*. **24**, 24-29.

Haag, J., O'hUigin, C., Overath, P. (1998). The molecular phylogeny of trypanosomes: evidence for an early divergence of the Salivaria. *Molecular & Biochemical Parasitology*. **91**, 37-49.

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

Hammarström, L., Marcotte, H. (2015). Passive immunization: toward magic bullets, In: Mucosal Immunology (Mestecky, J., Strober, W., Russell, M. W., Cheroutre, H., Lambrecht, B. N., Kelsall, B. L. (Eds.)): Oxford, Academic Press.

Hawlisch, H., zu Vilsendorf, A. M., Bautsch, W., Klos, A., Köhl, J. (2000). Guinea pig C3 specific rabbit single chain Fv antibodies from bone marrow, spleen and blood derived phage libraries. *Journal of Immunological Methods*. **236**, 117-131.

He, R., Drury, G. E., Rotari, V. I., Gordon, A., Willer, M., Farzaneh, T., Woltering, E. J., Gallois, P. (2008). Metacaspase-8 modulates programmed cell death induced by ultraviolet light and H_2O_2 in *Arabidopsis*. *Journal of Biological Chemistry*. **283**, 774-783.

Helms, M. J., Ambit, A., Appleton, P., Tetley, L., Coombs, G. H., Mottram, J. C. (2006). Bloodstream form *Trypanosoma brucei* depend on multiple metacaspases associated with RAB-11 positive endosomes. *Journal of Cell Science*. **119**, 1105-1117.

Hemerly, J., Oliveira, J., Del Nery, E., Morty, R. E., Andrews, N. W., Juliano, M. A., Juliano, L. (2003). Subsite specificity (S3, S2, S1 , S2 and S3) of oligopeptidase B from *Trypanosoma cruzi* and *Trypanosoma brucei* using fluorescent quenched peptides: comparative study and identification of specific carboxypeptidase activity. *Biochemical Journal*. **373**, 933–939.

Hertveldt, K., Belien, T., Volckaert, G. (2009). General M13 phage display: M13 phage display in identification and characterization of protein-protein interactions. *Methods in Molecular Biology*. **502**, 321-339.

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

Hide, G., Gray, A., Harrison, C. M., Tait, A. (1989). Identification of an epidermal growth factor receptor homologue in trypanosomes. *Molecular & Biochemical Parasitology*. **36**, 51-59.

Hirumi, H., Hirumi, K. (1991). *In vitro* cultivation of *Trypanosoma congolense* bloodstream forms in the absence of feeder cell layers. *Parasitology*. **102(Suppl. 2)**, 225-236.

Hoet, R. M., Cohen, E. H., Kent, R. B., Rookey, K., Schoonbroodt, S., Hogan, S., Rem, L., Frans, N., Daukandt, M., Pieters, H., van Hegelsom, R., Neer, N. C., Nastri, H. G., Rondon, I. J., Leeds, J. A., Hufton, S. E., Huang, L., Kashin, I., Devlin, M., Kuang, G., Steukers, M., Viswanathan, M., Nixon, A. E., Sexton, D. J., Hoogenboom, H. R., Ladner, R. C. (2005). Generation of high-affinity human antibodies by combining donor-derived and synthetic complementarity-determining-region diversity. *Nature Biotechnology*. **23**, 344-348.

Holmes, P. H., Eisler, M. C., Geerts, S. (2004). Current chemotherapy of animal trypanosomiasis, In: The Trypanosomiases, (Maudlin, I., Holmes, P., Miles, M. (Eds.)): Wallingford, CABI Publishing.

Holt, L. J., Herring, C., Jespers, L. S., Woolven, B. P., Tomlinson, I. M. (2003). Domain antibodies: proteins for therapy. *Trends in Biotechnology*. **21**, 484-490.

Hölzmuller, P., Grebaut, P., Brizard, J. P., Berthier, D., Chantal, I., Bossard, G., Bucheton, B., Vezilier, F., Chuchana, P., Bras-Gonçalves, R., Lemesre, J.-L., Vincendeau, P., Cuny, G., Frutos, R., Birond, D. G. (2008). Pathogeno-proteomics. *Annals of the New York Academy of Sciences*. **1149**, 66-70.

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. (2005). Selecting and screening recombinant antibody libraries. *Nature Biotechnology*. **23**, 1105-1116.

Horn, D., McCulloch, R. (2010). Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. *Current Opinion in Microbiology*. **13**, 700-705.

Hotez, P. J., Fenwick, A., Savioli, L., Molyneux, D. H. (2009). Rescuing the bottom billion through control of neglected tropical diseases. *Lancet*. **373**, 1570-1575.

Hovel-Miner, G., Mugnier, M. R., Goldwater, B., Cross, G. A. M., Papavasiliou, F. N. (2016). A conserved DNA repeat promotes selection of a diverse repertoire of *Trypanosoma brucei* surface antigens from the genomic archive. *PLoS Genetics*. **12**, e1005994.

Huang, S. Y., Zou, X. (2010). Advances and challenges in protein-ligand docking. *International Journal of Molecular Sciences*. **11**, 3016-3034.

Hudson, K. M., Terry, R. J. (1979). Immunodepression and the course of infection of a chronic *Trypanosoma brucei* infection in mice. *Parasite Immunology*. **1**, 317-326.

Huson, L., (2006). Antibody-mediated inhibition of proteases of African trypasomes. University of KwaZulu-Natal, Pietermaritzburg. PhD.

Hust, M., Dübel, S. (2004). Mating antibody phage display with proteomics. *Trends in Biotechnology*. **22**, 8-14.

Hust, M., Dübel, S. (2005). Phage display vectors for the in vitro generation of human antibody fragments. *Methods in Molecular Biology*. **295**, 71-96.

Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoler, R. E., Haber, E., Crea, R., Oppermann, H. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA*. **85**, 5879-5883.

Isaac, C., Ciosi, M., Hamilton, A., Scullion, K. M., Dede, P., Igbinosa, I. B., Nmorsi, O. P. G., Masiga, D., Turner, C. M. R. (2016). Molecular identification of different trypanosome species and subspecies in tsetse flies of northern Nigeria. *Parasites & Vectors*. **9**, 301.

Jackson, A. P., Goyard, S., Xia, D., Foth, B. J., Sanders, M., Wastling, J. M., Minoprio, P., Berriman, M. (2015). Global gene expression profiling through the complete life cycle of *Trypanosoma vivax*. *PLoS Neglected Tropical Diseases*. **9**, e0003975.

Jacobs, R. T., Nare, B., Wring, S. A., Orr, M. D., Chen, D., Sligar, J. M., Jenks, M. X., Noe, R. A., Bowling, T. S., Mercer, L. T., Rewerts, C., Gaukel, E., Owens, J., Parham, R., Randolph, R., Beaudet, B., Bacchi, C. J., Yarlett, N., Plattner, J. J., Freund, Y., Ding, C., Akama, T., Zhang, Y. K., Brun, R., Kaiser, M., Scandale, I., Don, R. (2011). SCYX-7158, an orally-active benzoxaborole for the treatment of stage 2 human African trypanosomiasis. *PLoS Neglected Tropical Diseases*. **5**, e1151.

Jain, V., Saravanan, P., Arvind, A., Mohan, C. G. (2011). First pharmacophore model of CCR3 receptor antagonists and its homology model-assisted, stepwise virtual screening. *Chemical biology & drug design*. **77**, 373-387.

Jamonneau, V., Camara, O., Ilboudo, H., Peylhard, M., Koffi, M., Sakande, H., N'Dri, L., Sanou, D., Dama, E., Camara, M., Lejon, V. (2015). Accuracy of individual rapid tests for serodiagnosis of *gambiense s*leeping sickness in west Africa. *PLoS Neglected Tropical Diseases*. **9**, e0003480.

Janeway Jnr, C. A., Medzhitov, R. (2002). Innate immune recognition. *Annual Review of Immunology*. **20**, 197-216.

Jetton, N., Rothberg, K. G., Hubbard, J. G., Wise, J., Li, Y., Ball, H. L., Ruben, L. (2009). The cell cycle as a therapeutic target against *Trypanosoma brucei*: Hesperadin inhibits Aurora kinase-1 and blocks mitotic progression in bloodstream forms. *Molecular Microbiology*. **72**, 442-458.

Jones, D. T., Taylor, W. R., Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences*. **8**, 275-282.

Jones, N. G., Nietlispach, D., Sharma, R., Burke, D. F., Eyres, I., Mues, M., Mott, H. R., Carrington, M. (2008). Structure of a glycosylphosphatidylinositol-anchored domain from a trypanosome variant surface glycoprotein. *Journal of Biological Chemistry*. **283**, 3584-3593.

Jones, T. W., Dávila, A. M. (2001). *Trypanosoma vivax*-out of Africa. *Trends in Parasitology*. **17**, 99-101.

Kaczanowski, S., Sajid, M., Reece, S. E. (2011). Evolution of apoptosis-like programmed cell death in unicellular protozoan parasites. *Parasites & Vectors*. **4**, 44.

Kajunguri, D., Hargrove, J. W., Ouifki, R., Mugisha, J. Y. T., Coleman, P. G., Welburn, S. C. (2014). Modelling the use of insecticide-treated cattle to control tsetse and *Trypanosoma*

brucei rhodesiense in a multi-host population. *Bulletin of Mathematical Biology*. **76**, 673-696.

Kalyaanamoorthy, S., Chen, Y. P. (2011). Structure-based drug design to augment hit discovery. *Drug Discovery Today*. **16**, 831-839.

Kanatani, A., Masuda, T., Shimoda, T., Misoka, F., Lin, X. S., Yoshimoto, T., Tsuru, D. (1991). Protease II from *Escherichia coli*: sequencing and expression of the enzyme gene and characterization of the expressed enzyme. *Journal of Biochemistry*. **110**, 315-320.

Kangethe, R. T., (2011). Gene disruption of TcoCATL (congopain) and oligopeptidase B, pathogenic factors of African trypanosomes. University of KwaZulu-Natal, Pietermaritzburg. PhD.

Kapetanovic, I. M. (2008). Computer-aided drug discovery and development (CADDD): In silico-chemico-biological approach. *Chemico-Biological Interactions*. **171**, 165-176.

Kashiwazaki, Y., Snowden, K., Smith, D. H., Hommel, M. (1994). A multiple antigen detection dipstick colloidal dye immunoassay for the field diagnosis of trypanosome infections in cattle. *Veterinary Parasitology*. **55**, 57-69.

Kaushik, R. S., Uzonna, J. E., Zhang, Y., Gordon, J. R., Tabel, H. (2000). Innate resistance to experimental African trypanosomiasis: differences in cytokine (TNF-α, IL-6, IL-10 and IL-12) production by bone marrow-derived macrophages from resistant and susceptible mice. *Cytokine*. **12**, 1024-1034.

Kay, B. K., Williamson, M. P., Sudol, M. (2000). The importance of being proline: the interaction of proline-rich motifs in signalling proteins with their cognate domains. *FASEB Journal*. **14**, 231-241.

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

Kgori, P. M., Modo, S., Torr, S. J. (2006). The use of aerial spraying to eliminate tsetse from the Okavango Delta of Botswana. *Acta Tropica*. **99**, 184-199.

Khare, S., Nagle, A. S., Biggart, A., Lai, Y. H., Liang, F., Davis, L. C., Barnes, W. S., Mathison, C. J. N., Myburgh, E., Gao, M., Gillespie, R. J., Liu, X., Tan, J. L., Stinson, M., Rivera, I. C., Ballard, J., Yeh, V., Groessl, T., Federe, G., Koh, H. X. Y., Venable, J. D., Bursulaya, B., Shapiro, M., Mishra, P. K., Spraggon, G., Brock, A., Mottram, J. C., Buckner, F. S., Rao, S. P. S., Wen, B. G., Walker, J. R., Tuntland, T., Molteni, V., Glynne, R. J., Supek, F. (2016). Proteasome inhibition for treatment of leishmaniasis, Chagas disease and sleeping sickness. *Nature*. **537**, 229-233.

Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B. A., Wang, J., Yu, B., Zhang, J., Bryant, S. H. (2016). PubChem substance and compound databases. *Nucleic Acids Research*. **44**, D1202-D1213.

Kipriyanov, S. M., Moldenhauer, G., Little, M. (1997). High level production of soluble single chain antibodies in small-scale *Escherichia coli* cultures. *Journal of Immunological Methods*. **200**, 69-77.

Kitchen, D. B., Decornez, H., Furr, J. R., Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery*. **3**, 935-949.

Klebe, G. (2006). Virtual ligand screening: Strategies, perspectives and limitations. *Drug Discovery Today*. **11**, 580-594.

Klemenčič, M., Novinec, M., Dolinar, M. (2015). Orthocaspases are proteolytically active prokaryotic caspase homologues: the case of *Microcystis aeruginosa*. *Molecular Microbiology*. **98**, 142-150.

Klingbeil, M. M., Englund, P. T. (2004). Closing the gaps in kinetoplast DNA network replication. *Proceedings of the National Academy of Sciences of the United States of America*. **101**, 4333-4334.

Kobayakawa, T., Louis, J., Izui, S., Lambert, P. H. (1979). Autoimmune response to DNA, red blood cells, and thrombocyte antigens in association with polyclonal antibody synthesis during experimental African trypanosomiasis. *Journal of Immunology*. **122**, 296-301.

Köhler, G., Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*. **256**, 495-497.

Koolpe, M., Burgess, R., Dail, M., Pasquale, E. B. (2005). EphB receptor-binding peptides identified by phage display enable design of an antagonist with ephrin-like affinity. *Journal of Biological Chemistry*. **280**, 17301-17311.

Kosec, G., Alvarez, V. E., Agüero, F., Sánchez, D., Doliinar, M., Turk, B., Cazzulo, J. J. (2006). Metacaspases of *Trypanosoma cruzi*: Possible candidates for programmed cell death mediators. *Molecular & Biochemical Parasitology*. **145**, 18-28.

Krause, R. G. E., (2016). Developing antibodies against *Plasmodium* lactate dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase and phosphoethanolamine-N-methyltransferase. University of KwaZulu-Natal, Pietermaritzburg. PhD.

Krawczyk, K., Liu, X., Baker, T., Shi, J., Deane, C. M. (2014). Improving B-cell epitope prediction and its application to global antibody-antigen docking. *Bioinformatics*. **30**, 2288-2294.

Kuhn, P., Fuhner, V., Unkauf, T., Moreira, G. M., 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.

Kulkarni, R. G., Srivani, P., Achaiah, G., Sastry, G. N. (2007). Strategies to design pyrazolyl urea derivatives for p38 kinase inhibition: a molecular modeling study. *Journal of Computer-Aided Molecular Design*. **21**, 155-166.

Kumar, S., Stecher, G., Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology & Evolution*. **33**, 1870-1874.

Kuzoe, F. A. S. (1993). Current situation of African trypanosomiasis. *Acta Tropica*. **54**, 153-162.

La Greca, F., Magez, S. (2011). Vaccination against trypanosomiasis, can it be done or is the trypanosome truly the ultimate immune destroyer and escape artist? *Human Vaccines*. **7**, 1225-1233.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. **277**, 680-685.

Lam, E., Zhang, Y. (2012). Regulating the reapers: activating metacaspases for programmed cell death. *Trends in Plant Science*. **17**, 487-494.

Lamkanfi, M., Festjens, N., Declercq, W., Van den Berghe, T., Vandenabeele, P. (2007). Caspases in cell survival, proliferation and differentiation. *Cell Death & Differentiation*. **14**, 44-55.

Langousis, G., Hill, K. L. (2014). Motility and more: the flagellum of *Trypanosoma brucei*. *Nature Reviews Microbiology*. **12**, 505-518.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*. **23**, 2947-2948.

Laverrière, M., Cazzulo, J. J., Alvarez, V. E. (2012). Antagonic activities of *Trypanosoma cruzi* metacaspases affect the balance between cell proliferation, death and differentiation. *Cell Death & Differentiation*. **19**, 1358-1369.

Lee, N., Gannavaram, S., Selvapandiyan, A., Debrabant, A. (2007). Characterization of metacaspases with trypsin-like activity and their putative role in programmed cell death in the protozoan parasite *Leishmania*. *Eukaryotic Cell*. **6**, 1745-1757.

Lee, R. E., Puente, L. G., Kaem, M., Megeney, L. A. (2008). A non-death role of the yeast metacaspase: Yca1p alters cell cycle dynamics. *PLoS ONE*. **3**, e2956.

Lee, R. E., Brunette, S., Puente, L. G., Megeney, L. A. (2010). Metacaspase Yca1 is required for clearance of insoluble protein aggregates. *Proceedings of the National Academy of Sciences*. **107**, 13348-13353.

Lee, Y., Kim, S., Kim, J. Y., Arooj, M., Kim, S., Hwang, S., Kim, B., Park, K. H., Lee, K. W. (2014). Binding mode analyses and pharmacophore model development for stilbene derivatives as a novel and competitive class of α-glucosidase inhibitors. *PLoS ONE*. **9**, e85827.

Lejon, V., Ngoyi, D. M., Boelaert, M., Büscher, P. (2010). A CATT negative result after treatment for human African trypanosomiasis is not indication for cure. *PLoS Neglected Tropical Diseases*. **4**, e590.

Lenardo, M. J., Rice-Ficht, A. C., Kelly, G., Esser, K. M., Donelson, J. E. (1984). Characterization of the genes specifying two metacyclic variable antigen types in *Trypanosoma brucei rhodesiense*. *Proceedings of the National Academy of Sciences of the United States of America*. **81**, 6642-6646.

Levenhagen, M. A., de Almeida Araújo Santos, F., Fujimura, P. T., Paula, A., Costa-Cruz, J. M., Goulart, L. R. (2015). Structural and functional characterization of a novel scFv anti-HSP60 of *Strongyloides* sp. *Scientific Reports*. **5**, 10447.

Li, M., Wang, H., Liu, J. X., Hao, P., Ma, L., Liu, Q. (2016). The apoptotic role of metacaspase in *Toxoplasma gondii*. *Frontiers in Microbiology*. **6**, 1560.

Li, Y., Kilpatrick, J., Whitelam, G. C. (2000). Sheep monoclonal antibody fragments generated using a phage display system. *Journal of Immunological Methods*. **236**, 133-146.

Li, Z., Umeyama, T., Wang, C. C. (2009). The Aurora kinase in *Trypanosoma brucei* plays distinctive roles in metaphase-anaphase transition and cytokinetic initiation. *PLoS Pathogens*. **5**, e1000575.

Liu, M. K., Pearson, T. W., Sayer, P. D., Gould, S. S., Waitumbi, J. N., Njogu, A. R. (1988). Serodiagnosis of African sleeping sickness in vervet monkeys by detection of parasite antigens. *Acta Tropica*. **45**, 321-330.

Lounnas, V., Ritschel, T., Kelder, J., McGuire, R., Bywater, R. P., Foloppe, N. (2013). Current progress in structure-based rational drug design marks a new mindset in drug discovery. *Computational & Structural Biotechnology*. **5**, e201302011.

Lucius, R., Los-Frank, B., Lane, R. P., Poulin, R., Roberts, C. W., Grencis, R. K. (2017). The biology of parasites. Wiley-VCH, Weinheim, Germany.

Luckins, A. G., Mehlitz, D. (1976). Observations on serum immunoglobulin levels in cattle infected with *Trypanosoma brucei, T. vivax* and *T. congolense*. *Annals of Tropical Medicine & Parasitology*. **70**, 479-480.

Lumbala, C., Bessell, P. R., Lutumba, P., Baloji, S., Biéler, S., Ndung'u, J. M. (2017). Performance of the SD BIOLINE® HAT rapid test in various diagnostic algorithms for *gambiense* human African trypanosomiasis in the Democratic Republic of the Congo. *PLoS ONE*. **12**, e0180555.

Luquetti, A. O., Ponce, C., Ponce, E., Esfandiari, J., Schijman, A., Revollo, S., Anez, N., Zingales, B., Ramgel-Aldao, R., Gonzalez, A., Levin, M. J., Umezawa, E. S., Franco da Silveira, J. (2003). Chagas' disease diagnosis: a multicentric evaluation of Chagas Stat-Pak, a rapid immunochromatographic assay with recombinant proteins of *Trypanosoma cruzi*. *Diagnostic Microbiology & Infectious Disease*. **46**, 265-271.

Lutje, V., Seixas, J., Kennedy, A. (2010). Chemotherapy for second-stage Human African trypanosomiasis. *The Cochrane Database of Systematic Reviews*. **4**, CD006201.

Mabbott, N. A., Sutherland, I. A., Sternberg, J. M. (1994). *Trypanosoma brucei* is protected from the cytostatic effects of nitric oxide under *in vivo* conditions. *Parasitology Research*. **80**, 687-690.

Maccari, G., Jaeger, T., Moraca, F., Biava, M., Flohe, L., Botta, M. (2011). A fast virtual screening approach to identify structurally diverse inhibitors of trypanothione reductase. *Bioorganic & Medicinal Chemistry Letters*. **21**, 5255-5258.

MacGregor, P., Savill, N. J., Hall, D., Matthews, K. R. (2011). Transmission stages dominate trypanosome within-host dynamics during chronic infections. *Cell Host & Microbe*. **9**, 310-318.

MacGregor, P., Szoor, B., Savill, N. J., Matthews, K. R. (2012). Trypanosomal immune evasion, chronicity and transmission: an elegant balancing act. *Nature Reviews Microbiology*. **10**, 431-438.

Machado, M. F. M., Marcondes, M. F., Juliano, M. A., McLuskey, K., Mottram, J. C., Moss, C. X., Juliano, L., Oliveira, V. (2013). Substrate specificity and the effect of calcium on *Trypanosoma brucei* metacaspase 2. *FEBS Journals*. **280**, 2608-2621.

Madeo, F. (2002). A caspase-related protease regulates apoptosis in yeast. *Molecular Cell*. **9**, 911-917.

Madeo, F., Herker, E., Maldener, C. (2002). A caspase-related protease regulates apoptosis in yeast. *Molecular Cell*. **9**, 911-917.

Magnus, E., Vervoort, T., Van Meirvenne, N. (1978). A card-agglutination test with stained trypanosomes (C.A.T.T.) for the serological diagnosis of *T. b. gambiense* trypanosomiasis. *Annales de la Societe Belge de Medecine Tropicale*. **58**, 169-176.

Mandal, S., Moudgil, M. N., Mandal, S. K. (2009). Rational drug design. *European Journal of Pharmacology*. **625**, 90-100.

Mansfield, J. M., Paulnock, D. M. (2005). Regulation of innate and acquired immunity in African trypanosomiasis. *Parasite Immunology*. **27**, 361-371.

Marcello, L., Barry, J. D. (2007). Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Research*. **17**, 1344-1352.

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H.,

Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., Geer, L. Y., Bryant, S. H. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research*. **45**, D200-D203.

Marston, F. A. O. (1986). The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochemical Journal*. **240**, 1-12.

Masake, R. A., Nantulya, V. M. (1991). Sensitivity of an antigen detection enzyme immunoassay for diagnosis of *Trypanosoma congolense* infections in goats and cattle. *Journal of Parasitology*. **77**, 231-236.

Maser, P., Wittlin, S., Rottmann, M., Wenzler, T., Kaiser, M., Brun, R. (2012). Antiparasitic agents: new drugs on the horizon. *Current Opinion in Pharmacology*. **12**, 562-566.

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

Matthews, K. R. (2015). 25 years of African trypanosome research: From description to molecular dissection and new drug discovery. *Molecular & Biochemical Parasitology*. **200**, 30-40.

Maudlin, I., Welburn, S. C. (1994). Maturation of trypanosome infections in tsetse. *Experimental Parasitology*. **79**, 202-205.

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

McCafferty, J. (1996). Phage display: factors affecting panning efficiency, In: Phage display of peptides and proteins: A laboratory manual, (Kay, B. K., Winter, J., McCafferty, J. (Eds.)): San Diego, California, Academic Press Inc.

McGrath, M. E., Gillmor, S. A., Fletterick, R. J. (1995). Ecotin: lessons on survival in a protease-filled world. *Protein Science*. **4**, 141-148.

McKerrow, J. H., Caffrey, C. R., Kelley, L. A., Loke, P., Sajid, M. (2006). Proteases in parasitic diseases. *Annual Review of Pathology*. **1**, 497-536.

McKerrow, J. H., Rosenthal, P. J., Swenerton, R., Doyle, P. (2008). Development of protease inhibitors for protozoan infections. *Current Opinion in Infectious Diseases*. **21**, 668-672.

McLuskey, K., Paterson, N. G., Bland, N. D., Isaacs, N. W., Mottram, J. C. (2010). Crystal stucture of *Leishmania major* oligopeptidase B gives insight into the enzymatic properties of a trypanosomatid virulence factor. *Journal of Biological Chemistry*. **285**, 39249-39259.

McLuskey, K., Rudolf, J., Proto, W. R., Isaacs, N. W., Coombs, G. H., Moss, C. X., Mottram, J. C. (2012). Crystal structure of a *Trypanosoma brucei* metacaspase. *PNAS*. **109**, 7469-7474.

McLuskey, K., Mottram, J. C. (2015). Comparative structural analysis of the caspase family with other clan CD cysteine peptidases. *Biochemical Journal*. **466**, 219-232.

Medina-Acosta, E., Cross, G. A. M. (1993). Rapid isolation of DNA from trypanosomatid protozoa using a simple 'mini-prep' procedure. *Molecular & Biochemical Parasitology*. **59**, 327-329.

Mendez-Lucio, O., Romo-Mancillas, A., Medina-Franco, J. L., Castillo, R. (2012). Computational study on the inhibition mechanism of cruzain by nitrile-containing molecules. *Journal of molecular graphics & modelling*. **35**, 28-35.

Mendoza-Palomares, C., Biteau, N., Giroud, C., Coustou, V., Coetzer, T. H. 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.

Meng, X. Y., Zhang, H. X., Mezei, M., Cui, M. (2011). Molecular docking: A powerful approach for structure-based drug discovery. *Current Computer-Aided Drug Design*. **7**, 146-157.

Meslin, B., Barnadas, C., Boni, V., Latour, C., De Monbrison, F., Kaiser, K., Picota, S. (2007). Features of apoptosis in *Plasmodium falciparum* erythrocytic stage through a putative role of PfMCA1 metacaspase-like protein. *Journal of Infectious Diseases*. **195**, 1852-1859.

Meslin, B., Zalila, H., Fasel, N., Picot, S., Bienvenu, A.-L. (2011). Are protozoan metacaspases potential parasite killers? *Parasites & Vectors*. **4**, 26-31.

Mitashi, P., Hasker, E., Lejon, V., Kande, V., Muyembe, J., Lutumba, P., Boelaert, M. (2012). Human African trypanosomiasis diagnosis in first-line health services of endemic countries, a systematic review. *PLoS Neglected Tropical Diseases*. **6**, e1919.

Morrison, L. J., Marcello, L., McCulloch, R. (2009). Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. *Cellular Microbiology*. **11**, 1724-1734.

Morty, R. E., Troeberg, L., Pike, R. N., Jones, R., Nickel, P., Lonsdale-Eccles, J. D., Coetzer, T. H. T. (1998). A trypanosome oligopeptidase as a target for the trypanocidal agents pentamidine, diminazene and suramin. *FEBS Letters*. **433**, 251-256.

Morty, R. E., Lonsdale-Eccles, J. D., Morehead, J., Caler, E. V., Mentele, R., Auerswald, E. A., Coetzer, T. H. T., Andrews, N. W., Burleigh, B. A. (1999). Oligopeptidase B from *Trypanosoma brucei*, a new member of an emerging subgroup of serine oligopeptidases. *Journal of Biological Chemistry*. **274**, 26149-26156.

Morty, R. E., Troeberg, L., Powers, J. C., Ono, S., Lonsdale-Eccles, J. D., Coetzer, T. H. T. (2000). Characterisation of the anti-trypanosomal activity of peptidyl α-aminoalkyl phosphate diphenyl esters. *Biochemical Pharmacology*. **60**, 1497-1504.

Morty, R. E., Lonsdale-Eccles, J. D., Mentele, R., Auerswald, E. A., Coetzer, T. H. T. (2001). Trypanosome-derived oligopeptidase B is released into the plasma of infected rodents, where it persists and retains full catalytic activity. *Infection & Immunity*. **69**, 2757-2761.

Morty, R. E., Pellé, R., Vadász, I., Uzcanga, G. L., Seeger, W., Bubis, J. (2005a). Oligopeptidase B from *Trypanosoma evansi*. A parasite peptidase that inactivates atrial natriuretic factor in the bloodstream of infected hosts. *Journal of Biological Chemistry*. **280**, 10925-10937.

Morty, R. E., Shih, A. Y., Fülöp, V., Andrews, N. W. (2005b). Identification of the reactive cysteine residues in oligopeptidase B from *Trypanosoma brucei*. *FEBS Letters*. **579**, 2191-2196.

Moss, C. X., Westrop, G. D., Juliano, L., Coombs, G. H., Mottram, J. C. (2007). Metacaspase 2 of *Trypanosoma brucei* is a calcium-dependent cysteine peptidase without processing. *FEBS Letters*. **581**, 5635-5639.

Moss, C. X., Brown, E., Hamilton, A., Van der Veken, P., Augustyns, K., Mottram, J. C. (2015). An essential signal peptide peptidase identified in an RNAi screen of serine peptidases of *Trypanosoma brucei*. *PLoS ONE*. **10**, e0123241.

Mosser, D. M., Roberts, J. F. (1982). *Trypanosoma brucei*: recognition *in vitro* of two developmental forms by murine macrophages. *Experimental Parasitology*. **54**, 310-316.

Mottram, J. C., Helms, M. J., Coombs, G. H., Sajid, M. (2003). Clan CD cysteine peptidases of parasitic protozoa. *Trends in Parasitology*. **19**, 182-187.

Mueller, G. A., Ankney, J. A., Glesner, J., Khurana, T., Edwards, L. L., Pedersen, L. C., Perera, L., Slater, J. E., Pomes, A., London, R. E. (2014). Characterization of an anti-Bla g 1 scFv: Epitope mapping and cross-reactivity. *Molecular Immunology*. **59**, 200-207.

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**, e1005784.

Muhanguzi, D., Picozzi, K., Hattendorf, J., Thrusfield, M., Welburn, S. C., Kabasa, J. D., Waiswa, C. (2014). Improvements on restricted insecticide application protocol for control of human and animal African trypanosomiasis in Eastern Uganda. *PLoS Neglected Tropical Diseases*. **8**, e3284.

Muhanguzi, D., Okello, W. O., Kabasa, J. D., Waiswa, C., Welburn, S. C., Shaw, A. P. M. (2015). Cost analysis of options for management of African animal trypanosomiasis using interventions targeted at cattle in Tororo district; south-eastern Uganda. *Parasites & Vectors*. **8**, 387.

Munday, J. C., McLuskey, K., Brown, E., Coombs, G. H., Mottram, J. C. (2011). Oligopeptidase B deficient mutants of *Leishmania major*. *Molecular & Biochemical Parasitology*. **175**, 49-57.

Musinguzi, S. P., Suganuma, K., Asada, M., Laohasinnarong, D., Sivakumar, T., Yokoyama, N., Namangala, B., Sugimoto, C., Suzuki, Y., Xuan, X., Inoue, N. (2016). A PCR-based survey of animal African trypanosomosis and selected piroplasm parasites of cattle and goats in Zambia. *Journal of Veterinary Medical Science*. **78**, 1819-1824.

Nallamsetty, S., Waugh, D. S. (2007). A generic protocol for the expression and purification of recombinant proteins in *Escherichia coli* using a combinatorial His6-maltose binding protein fusion tag. *Nature Protocols*. **2**, 383-391.

Nantulya, V. M., Musoke, A. J., Rurangirwa, F. R., Saigar, N., Minja, S. H. (1987). Monoclonal antibodies that distinguish *Trypanosoma congolense, T. vivax* and *T. brucei*. *Parasite Immunology*. **9**, 421-431.

Nantulya, V. M., Lindqvist, K. J. (1989). Antigen-detection enzyme immunoassays for the diagnosis of *Trypanosoma vivax, T. congolense* and *T. brucei* infections in cattle. *Tropical Medicine and Parasitology*. **40**, 267-272.

Ndao, M., Beaulieu, C., Black, W. C., Isabel, E., Vasquez-Camargo, F., Nath-Chowdhury, M., Masse, F., Mellon, C., Methot, N., Nicoll-Griffith, D. A. (2014). Reversible cysteine protease inhibitors show promise for a Chagas disease cure. *Antimicrobial Agents & Chemotherapy*. **58**, 1167-1178.

Ndung'u, J. M., Wright, N. G., Jennings, F. W., Murray, M. (1992). Changes in atrial natriuretic factor and plasma renin activity in dogs infected with *Trypanosoma brucei*. *Parasitology Research*. **78**, 553-556.

Neu, H. C., Heppel, L. A. (1965). The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. *Journal of Biological Chemistry*. **240**, 3685-3692.

Nguyen, T. T., Ruttayaporn, N., Goto, Y., Kawazu, S., Sakurai, T., Inoue, N. (2015). A TeGM6-4r antigen-based immunochromatographic test (ICT) for animal trypanosomosis. *Parasitology Research*. **114**, 4319-4325.

Nishikata, M. (1984). Trypsin-like protease from soybean seeds. Purification and some properties. *Journal of Biochemistry*. **95**, 1169-1177.

Nissim, A., Hoogenboom, H. R., Tomlinson, I. M., Flynn, G., Midgley, C., Lane, D., Winter, G. (1994). Antibody fragments from a 'single pot' phage display library as immunochemical reagents. *EMBO Journal*. **13**, 692-698.

Njiokou, F., Laveissiere, C., Simo, G., Nkinin, S., Grebaut, P., Cuny, G., Herder, S. (2006). Wild fauna as a probable animal reservoir for *Trypanosoma brucei gambiense* in Cameroon. *Infection, Genetics & Evolution*. **6**, 147-153.

Njiru, Z. K., Constantine, C. C., Guya, S., Crowther, J., Kiragu, J. M., Thompson, R. C., Davila, A. M. (2005). The use of ITS1 rDNA PCR in detecting pathogenic African trypanosomes. *Parasitology Research*. **95**, 186-192.

Njogu, P. M., Guantai, E. M., Pavadai, E., Chibale, K. (2016). Computer-aided drug discovery approaches against the tropical infectious diseases Malaria, Tuberculosis, Trypanosomiasis, and Leishmaniasis. *ACS infectious diseases*. **2**, 8-31.

Nuttall, S. D., Humberstone, K. S., Krishnan, U. V., Carmichael, J. A., Doughty, L., Hattarki, M., Coley, A. M., Casey, J. L., Anders, A. F., Foley, M., Irving, R. A., Hudson, P. J. (2004). Selection and affinity maturation of IgNAR variable domains targeting *Plasmodium falciparum* AMA1. *Proteins*. **55**, 187-197.

O'Brien, P. M., Aitken, R., O'Neil, B. W., Campo, M. S. (1999). Generation of native bovine mAbs by phage display. *Proceedings of the National Academy of Sciences USA*. **96**, 640-645.

Obishakin, E., Stijlemans, B., Santi-Rocca, J., Vandenberghe, I., Devreese, B., Muldermans, S., Bastin, P., Magez, S. (2014). Generation of a nanobody targeting the paraflagellar rod protein of Trypanosomes. *PLoS ONE*. **9**, e115893.

Odiit, M., Coleman, P. G., Liu, W. C., McDermott, J. J., Fevre, E. M., Welburn, S. C., Woolhouse, M. E. (2005). Quantifying the level of under-detection of *Trypanosoma brucei rhodesiense* sleeping sickness cases. *Tropical Medicine & International Health*. **10**, 840-849.

Odongo, S., Sterckx, Y. G. J., Stijlemans, B., Pillay, D., Baltz, T., Muyldermans, S., Magez, S. (2016). An anti-proteome nanobody library approach yields a specific immunoassay for *Trypanosoma congolense* diagnosis targeting glycosomal aldolase. *PLoS Neglected Tropical Diseases*. **10**, e00004420.

OIE. (2013). Manual of diagnostic tests and vaccines for terrestrial animals. World Organisation for Animal Health, Paris.

Ojha, M., Cattaneo, A., Hugh, S., Pawlowski, J., Cox, J. A. (2010). Structure, expression and function of *Allomyces arbuscula* CDP II (metacaspase) gene. *Gene*. **457**, 25-34.

Okello, W. O., Muhanguzi, D., MacLeod, E. T., Welburn, S. C., Waiswa, C., Shaw, A. P. (2015). Contribution of draft cattle to rural livelihoods in a district of southeastern Uganda endemic for bovine parasitic diseases: an economic evaluation. *Parasites & Vectors*. **8**, 1-9.

Opperman, P. A., Maree, F. F., van Wyngaardt, W., Vosloo, W., Theron, J. (2012). Mapping of antigenic determinants on a SAT2 foot-and-mouth disease virus using chicken single-chain antibody fragments. *Virus Research*. **167**, 370-379.

Osório, A., Madruga, C. R., Desquesnes, M., Soares, C. O., Ribeiro, L. R. R., Costa, S. C. G. (2008). *Trypanosoma* (Duttonella) *vivax*: its biology, epidemiology, pathogenesis, and introduction in the New World-a review. *Memórias do Instituto Oswaldo Cruz*. **103**, 1-13.

Pan, W., Ogunremi, O., Wei, G., Shi, M., Tabel, H. (2006). CR3 (CD11b/CD18) is the major macrophage receptor for IgM antibody-mediated phagocytosis of African trypanosomes:

diverse effect on subsequent synthesis of tumor necrosis factor alpha and nitric oxide. *Microbes & Infection*. **8**, 1209-1218.

Paquet, C., Ancelle, T., Gastellu-Etchegorry, M., Castilla, J., Harndt, I. (1992). Persistence of antibodies to *Trypanosoma brucei gambiense* after treatment of human trypanosomiasis in Uganda. *Lancet*. **340**, 250.

Pays, E., Lips, S., Nolan, D., Vanhamme, L., Perez-Morga, D. (2001). The VSG expression sites of *Trypanosoma brucei*: multipurpose tools for the adaptation of the parasite to mammalian hosts. *Molecular & Biochemical Parasitology*. **114**, 1-16.

Pays, E., Vanhamme, L., Perez-Morga, D. (2004). Antigenic variation in *Trypanosoma brucei*: facts, challenges and mysteries. *Current Opinion in Microbiology*. **7**, 369-374.

Pays, E. (2006). The variant surface glycoprotein as a tool for adaptation in African trypanosomes. *Microbes & Infection*. **8**, 930-937.

Pays, E., Vanhollebeke, B., Vanhamme, L., Paturiaux-Hanocq, F., Nolan, D. P., Pérez-Morga, D. (2006). The trypanolytic factor of human serum. *Nature Reviews Microbiology*. **4**, 477-486.

Peeling, R. W., Holmes, K. K., Mabey, D., Ronald, A. (2006). Rapid tests for sexually transmitted infections (STIs): the way forward. *Sexually transmitted infections*. **82**, 1-6.

Peña, I., Pilar Manzano, M., Cantizani, J., Kessler, A., Alonso-Padilla, J., Bardera, A. I., Alvarez, E., Colmenarejo, G., Cotillo, I., Roquero, I., de Dios-Anton, F., Barroso, V., Rodriguez, A., Gray, D. W., Navarro, M., Kumar, V., Sherstnev, A., Drewry, D. H., Brown, J. R., Fiandor, J. M., Julio Martin, J. (2015). New compound sets identified from high throughput phenotypic screening against three kinetoplastid parasites: an open resource. *Scientific Reports*. **5**, 8771.

Peña, M. S., Cabral, G. C., Fotoran, W. L., Perez, K. R., Stolf, B. S. (2017). Metacaspase-binding peptide inhibits heat shock-induced death in *Leishmania (L.) amazonensis*. *Cell death & disease*. **8**, e2645.

Philibert, P., Stoessel, A., Wang, W., Sibler, A. P., Bec, N., Larroque, C., Saven, J. G., Courtête, J., Weiss, E., Martineau, P. (2007). A focused antibody library for selecting scFvs expressed at high levels in the cytoplasm. *BMC Biotechnology*. **7**, 81.

Picozzi, K., Fèvre, E., Odiit, M., Carrington, M., Eisler, M. C., Maudlin, I., Welburn, S. C. (2005). Sleeping sickness in Uganda: a thin line between two fatal diseases. *British Medical Journal*. **331**, 1238-1241.

Pillay, D., Izotte, J., Fikru, R., Büscher, P., Mucache, H., Neves, L., Boulangé, A., Seck, M. T., Bouyer, J., Napier, G. B., Chevtzoff, C., Coustou, V., Baltz, T. (2013). *Trypanosoma vivax* GM6 antigen: a candidate antigen for diagnosis of animal African trypanosomiasis in cattle. *PLoS ONE*. **8**, e78565.

Pils, B., Schultz, J. (2004). Inactive enzyme-homologues find new function in regulatory processes. *Journal of Molecular Biology*. **340**, 399-404.

Pini, A., Bracci, L. (2000). Phage display of antibody fragments. *Current Protein & Peptide Science*. **1**, 155-169.

Pinto Dias, J. C. (2006). The treatment of Chagas disease (South American trypanosomiasis). *Annals of internal medicine*. **144**, 772-774.

Piszczek, E., Dudkiewicz, M., Mielecki, M. (2012). Biochemical and bioinformatic characterization of type II metacaspase protein (TaeMCAII) from wheat. *Plant Molecular Biology Reporter*. **30**, 1338-1347.

Pop, C., Salvesen, G. S. (2009). Human caspases: activation, specificity, and regulation. *Journal of Biological Chemistry*. **284**, 21777-21781.

Porter, R. R., Press, E. M. (1962). Immunochemistry. *Annual Review of Biochemistry*. **31**, 625-652.

Posthuma-Trumpie, G., Korf, J., van Amerongen, A. (2009). Lateral flow (immuno)assay: its strengths, weaknesses, opportunities and threats. A literature survey. *Analytical & Bioanalytical Chemistry*. **393**, 569-582.

Priotto, G., Kasparian, S., Mutombo, W., Ngouama, D., Ghorashian, S., Arnold, U., Ghabri, S., Baudin, E., Buard, V., Kazadi-Kyanza, S., Ilunga, M., Mutangala, W., Pohlig, G., Schmid, C., Karunakara, U., Torreele, E., Kande, V. (2009). Nifurtimox-eflornithine combination therapy for second-stage African *Trypanosoma brucei gambiense* trypanosomiasis: a multicentre, randomised, Phase III, non-inferiority trial. *Lancet*. **374**, 56-64.

Proto, W. R., Castanys-Munoz, E., Black, A., Tetley, L., Moss, C. X., Juliano, L., Coombs, G. H., Mottram, J. C. (2011). *Trypanosoma brucei* metacaspase 4 is a pseudopeptidase and a virulence factor. *Journal of Biological Chemistry*. **286**, 39914-39925.

Puratchikody, A., Sriram, D., Umamaheswari, A., Irfan, N. (2016). 3-D structural interactions and quantitative structural toxicity studies of tyrosine derivatives intended for safe potent inflammation treatment. *Chemistry Central Journal*. **10**, 24.

Radwanska, M., Guirnalda, P., De Trez, C., Ryffel, B., Black, S. J., Magez, S. (2008). Trypanosomiasis-induced B cell apoptosis results in loss of protective anti-parasite antibody responses and abolishment of vaccine-induced memory responses. *PLoS Pathogens*. **4**, e1000078.

Raina, P., Kaur, S. (2012). Knockdown of LdMC1 and Hsp70 by antisense oligonucleotides causes cell-cycle defects and programmed cell death in *Leishmania donovani*. *Molecular & Cellular Biochemistry*. **359**, 135-149.

Rakabe, M., (2008). Selection of chicken single-chain antibody fragments directed against recombinant VP7 of bluetongue virus. University of Pretoria, Pretoria. MSc.

Rakabe, M., van Wyngaardt, W., Fehrsen, J. (2011). Chicken single-chain antibody fragments directed against recombinant VP7 of bluetongue virus. *Food & Agricultural Immunology*. **22**, 283-295.

Rand, K. N. (1996). Crystal violet can be used to visualize DNA bands during gel electrophoresis and to improve cloning efficiency. *Technical Tips Online*. **1**, 23-23.

Rawlings, N. D., Barrett, A. J., Finn, R. D. (2016). Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*. **44**, D343-D350.

Rea, D., Fülöp, V. (2006). Structure-function properties of prolyl oligopeptidase family enzymes. *Cell Biochemistry & Biophysics*. **44**, 349-365.

Reithinger, R., Quinnell, R. J., Alexander, B., Davies, C. R. (2002). Rapid detection of *Leishmania infantum* infection in dogs: comparative study using an immunochromatographic dipstick test, enzyme-linked immunosorbent assay, and PCR. *Journal of Clinical Microbiology*. **40**, 2352-2356.

Renart, J., Reiser, J., Stark, G. R. (1979). Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure. *Proceedings of the National Academy of Sciences USA*. **76**, 3116-3120.

Reynaud, C., Anquez, V., Dahan, A., Weill, J. (1985). A single rearrangement event generates most of the chicken immunoglobulin light chain diversity. *Cell*. **40**, 283-291.

Reynaud, C., Anquez, V., Grimal, H., Weill, J. C. (1987). A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell*. **48**, 379-388.

Reynaud, C. A., Dahan, A., Anquez, V., Weill, J. C. (1989). Somatic hyperconversion diversifies the single Vh gene of the chicken with a high incidence in the D region. *Cell*. **59**, 171-183.

Reynolds, S. L., Fischer, K. (2015). Pseudoproteases: mechanisms and function. *Biochemical Journal*. **468**, 17-24.

Ridder, R., Schmitz, R., Legay, F., Gram, H. (1995). Generation of rabbit monoclonal antibody fragments from a combinatorial phage display library and their production in the yeast *Pichia pastoris*. *Nature Biotechnology*. **13**, 255-260.

Riordan, J. F. (1977). The role of metals in enzyme activity. *Annals of Clinical & Laboratory Science*. **7**, 119-129.

Robays, J., Bilengue, M. M., Van der Stuyft, P., Boelaert, M. (2004). The effectiveness of active population screening and treatment for sleeping sickness control in the Democratic Republic of Congo. *Tropical Medicine & International Health*. **9**, 542-550.

Roberts, R. J., Belfortm, M., Bestor, T., Bhagwat, A. S., Bickle, T. A., Bitinaite, B., Blumenthal, R. M., Degtyarev, S. K., Dryden, D. T. F., Dybvig, K., Firman, K., Gromova, E. S., Gumport, R. I., Halford, S. E., Hattman, S. H., Heitman, J., Hornby, D. P., Janulaitis, A., Jeltsch, A., Josephsen, J., Kiss, A., Klaenhammer, T. R., Kobayashi, I., Kong, H., KruÈger, D. H., Lacks, S., Marinus, M. G., Miyahara, M., Morgan, R. D., Murray, N. E., Nagaraja, V., Piekarowicz, A., Pingoud, A., Raleigh, E., Rao, D. N., Reich, N., Repin, V. E., Selker, E. U., Shaw, P.-C., Stein, D. C., Stoddard, B. L., Szybalski, W., Trautner, T. A., Van Etten, J. L., Vitor, J. M. B., Wilson, G. G., Xu, S.-Y. (2003). A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Research*. **31**, 1805-1812.

Robin, G., Sato, Y., Desplancq, D., Rochel, N., Weiss, E., Martineau, P. (2014). Restricted diversity of antigen binding residues of antibodies revealed by computational alanine scanning of 227 antibody–antigen complexes. *Journal of Molecular Biology*. **426**, 3729-3743.

Roditi, I., Liniger, M. (2002). Dressed for success: the surface coats of insect-borne protozoan parasites. *Trends in Microbiology*. **10**, 128-134.

Roditi, I., Lehane, M. J. (2008). Interactions between trypanosomes and tsetse flies. *Current Opinion in Microbiology*. **11**, 345-351.

Roger, F. L., Solano, P., Bouyer, J., Porphyre, V., Berthier, D., Peyre, M., Bonnet, P. (2017). Advocacy for identifying certain animal diseases as "neglected". *PLOS Neglected Tropical Diseases*. **11**, e0005843.

Rooney, B., Peining, T., Büscher, P., Rogé, S., Smales, C. M. (2015). Expression of *Trypanosoma brucei gambiense* antigens in *Leishmania tarentolae*. Potential for use in rapid serodiagnostic tests (RDTs). *PLoS Neglected Tropical Diseases*. **9**, e0004271.

Ross, G. A., Morris, G. M., Biggin, P. C. (2012). Rapid and accurate prediction and scoring of water molecules in protein binding sites. *PLoS One*. **7**, e32036.

Rotureau, B., Subota, I., Buisson, J., Bastin, P. (2012). A new asymmetric division contributes to the continuous production of infective trypanosomes in the tsetse fly *Development*. **139**, 1842-1850.

Rotureau, B., Van Den Abbeele, J. (2013). Through the dark continent: African trypanosome development in the tsetse fly. *Frontiers in Cellular & Infection Microbiology*. **3**, 53.

Roy, S., Bayly, C., Gareau, Y., Houtzager, V. M., Kargman, S., Keen, S. L., Rowland, K., Seiden, I. M., Thornberry, N. A., Nicholson, D. W. (2001). Maintenance of caspase-3 proenzyme dormancy by an intrinsic "safety catch" regulatory tripeptide. *Proceedings of the National Academy of Sciences USA*. **98**, 6132-6137.

Roy, S., Sharom, J. R., Houde, C., Loisel, T. P., Vaillancourt, J. P., Shao, W., Saleh, M., Nicholson, D. W. (2014). Confinement of caspase-12 proteolytic activity to autoprocessing. *Proceedings of the National Academy of Sciences*. **105**, 4133-4138.

Rozen, S., Skaletsky, H. (2000). Primer3 on the WWW for general users and for the biologist programmers. *Methods in Molecular Biology*. **132**, 365-386.

Saerens, D., Stijlemans, B., Baral, T. N., Nguyen Thi, G. T., Wernery, U., Magez, S., De Baetselier, P., Muyldermans, S., Conrath, K. (2008). Parallel selection of multiple anti-infectome nanobodies without access to purified antigens. *Journal of Immunological Methods*. **329**, 138-150.

Salas-Sarduy, E., Landaburu, L. U., Karpiak, J. X., Madauss, K. P., Cazzulo, J. J., Agüero, F., Alvarez, V. E. (2017). Novel scaffolds for inhibition of Cruzipain identified from highthroughput screening of anti-kinetoplastid chemical boxes. *Scientific Reports*. **7**, 12073.

Salum, L., Polikarpov, I., Andricopulo, A. D. (2008). Structure-based approach for the study of estrogen receptor binding affinity and subtype selectivity. *Journal of Chemical Information & Modeling*. **48**, 2243-2253.

Salvesen, G., Nagase, H. (1989). Inhibition of proteolytic enzymes, In: Proteolytic Enzymes: A Practical Approach, (Beyon, R. J., Bond, J. S. (Eds.)): Oxford, IRL Press.

Sambrook, J., Russell, D. W., Irwin, N. (2001). Molecular cloning: a laboratory manual, 3rd Edition. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.

Sapats, S. I., Heine, H. G., Trinidad, L., Gould, G. J., Foord, A. J., Doolan, S. G., Prowse, S., Ignjatovic, J. (2003). Generation of chicken single chain antibody variable fragments (scFv) that differentiate and neutralize infectious bursal disease virus (IBDV). *Archives of Virology*. **148**, 497-515.

Sblattero, D., Bradbury, A. (2000). Exploiting recombination in single bacteria to make large phage antibody libraries. *Nature Biotechnology*. **18**, 75-80.

Schirrmann, T., Meyer, T., Schütte, M., Frenzel, A., Hust, M. (2011). Phage display for the generation fo antibodies for proteome research, diagnostics and therapy. *Molecules*. **16**, 412-426.

Schlager, B., Straessle, A., Hafen, E. (2012). Use of anionic denaturing detergents to purify insoluble proteins after overexpression. *BMC Biotechnology*. **12**, 95.

Schmidt, T., Bergner, A., Schwede, T. (2014). Modelling three-dimensional protein structures for applications in drug design. *Drug Discovery Today*. **19**, 890-897.

Schwede, A., Jones, N., Engstler, M., Carrington, M. (2011). The VSG C-terminal domain is inaccessible to antibodies on live trypanosomes. *Molecular & Biochemical Parasitology*. **175**, 201-204.

Shao, W., Yeretssian, G., Doiron, K., Hussain, S. N., Saleh, M. (2007). The caspase-1 digestome identifies the glycolysis pathway as a target during infection and septic shock. *Journal of Biological Chemistry*. **282**, 36321-36329.

Shaw, A. P., Wint, G. R., Cecchi, G., Torr, S. J., Mattioli, R. C., Robinson, T. P. (2015). Mapping the benefit-cost ratios of interventions against bovine trypanosomosis in Eastern Africa. *Preventive Veterinary Medicine*. **122**, 406-416.

Shi, B., Wang, H., Guo, S., Xu, Y., Li, Z., Gu, J. (2007). Protein III-based single-chain antibody phage display using bacterial cells bearing an additional genome of a gene-III-lacking helper phage. *Biotechniques*. **42**, 760-765.

Shi, M., Wei, G., Pan, W., Tabel, H. (2004). *Trypanosoma congolense* infections: antibody-mediated phagocytosis by Kupffer cells. *Journal of Leukocyte Biology*. **76**, 399-405.

Shih, H. H., Tu, C., Cao, W., Klein, A., Ramsey, R., Fennell, B. J., Lambert, M., Ni Shuilleabhain, D., Autin, B., Kouranova, E., Laxmanan, S., Braithwaite, S., Wu, L., Ait-Zahra, M., Milici, A. J., Dumin, J. A., LaVallie, E. R., Arai, M., Corcoran, C., Paulsen, J. E., Gill, D., Cunningham, O., Bard, J., Mosyak, L., Finlay, W. J. (2012). An ultra-specific avian antibody to phosphorylated tau protein reveals a unique mechanism for phosphoepitope recognition. *Journal of Biological Chemistry*. **287**, 44425-44434.

Silva, A., Almeida, B., Sampaio-Marques, B., Reis, M. I. R., Ohlmeier, S., Rodrigues, F., do Vale, A., Ludovico, P. (2011). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is a specific substrate of yeast metacaspase. *Biochemica et Biophysica Acta*. **1813**, 2044-2049.

Silverman, J. M., Chan, S. K., Robinson, D. P., Dwyer, D. M., Nandan, D., Foster, L. J., Reiner, N. E. (2008). Proteomic analysis of the secretome of *Leishmania donovani*. *Genome Biology*. **9**, R35.

Silzel, J. W., Cercek, B., Dodson, C., Tsay, T., Obremski, R. J. (1998). Mass-sensing, multianalyte microarray immunoassay with imaging detection. *Clinical Chemistry*. **44**, 2036-2043.

Simarro, P. P., Cecchi, G., Paone, M., Franco, J. R., Diarra, A., Ruiz, J. A., Févre, E. M., Courtin, F., Mattioli, R. C., Jannin, J. (2010). The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. *International Journal of Health Geographics*. **9**, 57.

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

Simpson, A. G. B., Stevens, J. R., Lukeš, J. (2006). The evolution and diversity of kinetoplastid flagellates. *Trends in Parasitology*. **22**, 168-174.

Sippl, M. J. (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins*. **17**, 355-362.

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

Skinner-Adams, T. S., Sumanadasa, S. D., Fisher, G. M., Davis, R. A., Doolan, D. L., Andrews, K. T. (2016). Defining the targets of antiparasitic compounds. *Drug Discovery Today*. **21**, 725-739.

Skottrup, P. D., Leonard, P., Kaczmarek, J. Z., Veillard, F., Enghild, J. J., O'Kennedy, R., Sroka, A., Clausen, R. P., Potempa, J., Riise, E. (2011). Diagnostic evaluation of a nanobody with picomolar affinity toward the protease RgpB from *Porphyromonas gingivalis*. *Analytical Biochemistry*. **415**, 158-167.

Sliwoski, G., Kothiwale, S., Meiler, J., Lowe, E. W., Jr. (2013). Computational methods in drug discovery. *Pharmacological Reviews*. **66**, 334-395.

Smialowski, P., Martin-Galiano, A. J., Mikolajka, A., Girschick, T., Holak, T. A., Frishman, D. (2006). Protein solubility: sequence based prediction and experimental verification. *Bioinformatics*. **23**, 2536-2542.

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

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., Klenk, D. C. (1985). Measurement of protein using bicinchoninic acid. *Analytical Biochemistry*. **150**, 76-85.

Sousa, S. F., Ribeiro, A. J., Coimbra, J., Neves, R. P., Martins, S. A., Moorthy, N. S., Fernandes, P. A., Ramos, M. J. (2013). Protein-ligand docking in the new millennium-A retrospective of 10 years in the field. *Current Medicinal Chemistry*. **20**, 2296-2314.

Spitzer, R., Jain, A. N. (2012). Surflex-Dock: Docking benchmarks and real-world application. *Journal of Computer-Aided Molecular Design*. **26**, 687-699.

Sternberg, J. M., Mabbott, N. A. (1996). Nitric oxide-mediated suppression of T cell responses during *Trypanosoma brucei* infection: soluble trypanosome products and interferon-γ are synergistic inducers of nitric oxide synthase. *European Journal of Immunology*. **26**, 539-543.

Sternberg, J. M., Gierliński, M., Biéler, S., Ferguson, M. A. J., Ndung'u, J. M. (2014). Evaluation of the diagnostic accuracy of prototype rapid tests for human African trypanosomiasis. *PLoS Neglected Tropical Diseases*. **8**, e3373.

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 relevent drug target. *International Journal for Parasitology*. **42**, 481-488.

Stijlemans, B., Conrath, K., Cortez-Retamozo, V., Van Xong, H., Wyns, L., Senter, P., Revets, H., De Baetselier, P., Muyldermans, S., Magez, S. (2004). Efficient targeting of conserved cryptic epitopes of infectious agents by single domain antibodies. African trypanosomes as paradigm. *Journal of Biological Chemistry*. **279**, 1256-1261.

Stoevesandt, O., Taussig, M. J., He, M. (2009). Protein microarrays: high-throughput tools for proteomics. *Expert Review of Proteomics*. **6**, 145-157.

Streltsov, V. A., Varghese, J. N., Carmichael, J. A., Irving, R. A., Hudson, P. J., Nuttall, S. D. (2004). Structural evidence for evolution of shark Ig new antigen receptor variable domain antibodies from a cell-surface receptor. *Proceedings of the National Academy of Sciences USA*. **101**, 12444-12449.

Stuart, K., Brun, R., Croft, S. L., Fairlamb, A., Gürtler, R. E., McKerrow, J. H., Reed, S., Tarleton, R. (2008). Kinetoplastids: related protozoan pathogens, different diseases. *Journal of Clinical Investigation*. **118**, 1301-1310.

Sullivan, L., Wall, S. J., Carrington, M., Ferguson, M. A. J. (2013). Proteomic selection of immunodiagnostic antigens for human African trypanosomiasis and generation of a prototype lateral flow immunodiagnostic device. *PLoS Neglected Tropical Diseases*. **7**, e2087.

Sullivan, L., Fleming, J., Sastry, L., Mehlert, A., Wall, S. J., Ferguson, M. A. (2014). Identification of sVSG117 as an immunodiagnostic antigen and evaluation of a dual-antigen lateral flow test for the diagnosis of human African trypanosomiasis. *PLoS Neglected Tropical Diseases*. **8**, e2976.

Sun, J., Pons, J., Craik, C. S. (2003). Potent and selective inhibition of membrane-type serine protease 1 by human single-chain antibodies. *Biochemistry*. **42**, 892-900.

Sundström, J. F., Vaculova, A., Smertenko, A. P., Savenkov, E. I., Golovko, A., Minina, E., Tiwari, B. S., Rodriguez-Nieto, S., Zamyatnin, J., A. A., Välineva, T., Saarikettu, J., Frilander, M. J., Suarez, M. F., Zavialov, A., Ståhl, U., Hussey, P. J., Silvennoinen, O., Sundberg, E., Zhivotovsky, B., Bozhkov, P. V. (2009). Tudor staphylococcal nuclease is an evolutionarily conserved component of the programmed cell death degradome. *Nature Cell Biology*. **11**, 1347-1354.

Swallow, B. (2000). Impacts of trypanosomiasis on African agriculture. *PAAT Technical Scientific Series*. **2**, 52.

Szallies, A., Kubata, B. K., Duszenko, M. (2002). A metacaspase of *Trypanosoma brucei* causes loss of respiration competence and clonal death in the yeast *Saccharomyces cerevisiae*. *FEBS Letters*. **517**, 144-150.

Tait, S. W. G., Green, D. R. (2010). Mitochondria and cell death: outer membrane permeabilization and beyond. *Nature Reviews Molecular Cell Biology*. **11**, 621-632.

Tarral, A., Blesson, S., Mordt, O. V., Torreele, E., Sassella, D., Bray, M. A., Hovsepian, L., Evene, E., Gualano, V., Felices, M., Strub-Wourgaft, N. (2014). Determination of an optimal dosing regimen for fexinidazole, a novel oral drug for the treatment of human African trypanosomiasis: first-in-human studies. *Clinical pharmacokinetics*. **53**, 565-580.

Taylor, J. E., Rudenko, G. (2006). Switching trypanosome coats: what's in the wardrobe? *Trends in Genetics*. **22**, 614-620.

Tetaert, D., Soudan, B., Huet-Duvillier, G., Degand, P., Boersma, A. (1993). Unusual cleavage of peptidic hormones generated by trypanosome enzymes released in infested rat serum. *International Journal for Peptide & Protein Research*. **41**, 147-152.

Thomas, W. D., Golomb, M., Smith, G. P. (2010). Corruption of phage display libraries by target-unrelated clones: Diagnosis and countermeasures. *Analytical biochemistry*. **407**, 237-240.

Thompson, C. B., Neiman, P. E. (1987). Somatic diversification of the chicken light chain gene is limited to the rearranged light chain segment. *Cell*. **48**, 369-378.

Thumbi, S. M., McOdimba, F. A., Mosi, R. O., Jung'a, J. O. (2008). Comparative evaluation of three PCR base diagnostic assays for the detection of pathogenic trypanosomes in cattle blood. *Parasites & Vectors*. **1**, 46.

Tiberti, N., Hainard, A., Sanchez, J.-C. (2013). Translation of human African trypanosomiasis biomarkers towards field application. *Translational Proteomics*. **1**, 12-24.

Torr, S. J., Maudlin, I., Vale, G. A. (2007). Less is more: restricted application of insecticide to cattle to improve the cost and efficacy of tsetse control. *Medical & Veterinary Entomology*. **21**, 53-64.

Troeberg, L., Pike, R. N., Morty, R. E., Berry, R. K., Coetzer, T. H. T., 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.

Tsiatsiani, L., Van Breusegem, F., Gallois, P., Zavialov, A., Lamm, E., Bozhkov, P. V. (2011). Metacaspases. *Cell Death & Differentiation*. **18**, 1279-1288.

Tsuji, A., Yuasa, K., Matsuda, Y. (2004). Identification of oligopeptidase B in higher plants. Purification and characterization of oligopeptidase B from quiescent wheat embryo, *Triticum aestivum*. *Journal of Biochemistry*. **136**, 673-681.

Tu, X., Kumar, P., Li, Z., Wang, C. C. (2006). An aurora kinase homologue is involved in regulating both mitosis and cytokinesis in *Trypanosoma brucei*. *Journal of Biological Chemistry*. **281**, 9677-9687.

Turner, M. J. (1985). The biochemistry of the surface antigens of the African trypanosomes. *British Medical Bulletin*. **41**, 137-143.

Uhlén, M., Forsberg, G., Moks, T., Hartmanis, M., Nilsson, B. (1992). Fusion proteins in biotechnology. *Current Opinion in Biotechnology*. **3**, 363-369.

Uilenberg, G., Boyt, W. P. (1998). A field guide for the diagnosis, treatment and prevention of African animal trypanosomosis. *Food & Agriculture Organization of the United Nations*.

Uren, A. G., O'Rourke, K., Aravind, L., Pisabarro, T. M., Seshagiri, S., Koonin, V. E., Dixit, M. V. (2000). Identification of paracaspases and metacaspases: two ancient families of caspase-like proteins, one of which plays a key role in MALT lymphoma. *Molecular Cell*. **6**, 961-967.

Urwyler, S., Struder, E., Renggli, C. K., Roditi, I. (2007). A family of stage-specific alanine-rich proteins on the surface of epimastigote forms of *Trypanosoma brucei*. *Molecular Microbiology*. **63**, 218-228.

Urwyler, S. (2011). Allosteric modulation of family C G-protein-coupled receptors: From molecular insights to therapeutic perspectives. *Pharmacological Reviews*. **63**, 59-125.

Vale, G., Torr, S. (2004). Development of bait technology to control tsetse, In: The Trypanosomiases, (Maudlin, I., Holmes, P., Miles, M. (Eds.)): Wallingford, CABI Publishing.

Van den Bossche, P., Doran, M., Connor, R. J. (2000). An analysis of trypanocidal drug use in Eastern province of Zambia. *Veterinary Record*. **81**, 567-568.

Van den Bossche, P., Delespaux, V. (2011). Options for the control of tsetse-transmitted livestock trypanosomosis. An epidemiological perspective. *Veterinary Parasitology*. **181**, 37-42.

Van Meirvenne, N., Magnus, E., Büscher, P. (1995). Evaluation of variant specific trypanolysis tests for serodiagnosis of human infections with *Trypanosoma brucei gambiense*. *Acta Tropica*. **60**, 189-199.

van Wyngaardt, W., Du Plessis, D. H. (1998). Selection of an scFv phage antibody that recognizes bluetongue virus from a large synthetic library and its use in ELISAs to detect viral antigen and antibodies. *Onderstepoort Journal of Veterinary Research*. **65**, 125-131.

van Wyngaardt, W., Malatji, T., Mashau, C., Fehrsen, J., Jordaan, F., Miltiadou, D., du Plessis, D. (2004). A large semi-synthetic single chain Fv phage display library based on chicken immunoglobulin genes. *BMC Biotechnology*. **4**, 6.

van Wyngaardt, W., Mashau, C., Wright, I., Fehrsen, J. (2013). Serotype- and serogroup-specific detection of African horsesickness virus using phage displayed chicken scFvs for indirect double antibody sandwich ELISAs. *Journal of Veterinary Science*. **14**, 95-98.

Vandana, K., Singh, A. P., Singh, J., Sharma, R., Akhter, M., Mishra, P. K., Saxena, A. K., Dixit, R., Rathi, B., Katyal, A., Pandey, K. C. (2018). Biochemical characterization of unusual cysteine protease of *P. falciparum*, metacaspase-2 (MCA-2). *Molecular & Biochemical Parasitology*. **IN PRESS**.

Vanhamme, L., Poelvoorde, P., Pays, A., Tebabi, P., Xong, H. V., Pays, E. (2000). Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. *Molecular Microbiology*. **36**, 328-340.

Vanhamme, L., Pays, E., McCulloch, R., Barry, J. D. (2001). An update on antigenic variation in African trypanosomes. *Trends in Parasitology*. **17**, 338-343.

Vanhollebeke, B., Pays, E. (2010). The trypanolytic factor of human serum: many ways to enter the parasite, a single way to kill. *Molecular Microbiology*. **76**, 806-814.

Venkatachalam, C. M., Jiang, X., Oldfield, T., Waldman, M. (2003). LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of molecular graphics & modelling*. **21**, 289-307.

Ventura, S., Villaverde, A. (2006). Protein quality in bacterial inclusion bodies. *Trends in Biotechnology*. **24**, 179-185.

Vercammen, D., van de Cotte, B., De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K., Vandenberghe, I., Van Beeumen, J., Inze, D., Van Breusegem, F. (2004). Type II metacaspases Atmc4 and Atmc9 of *Arabidopsis thaliana* cleave substrates after arginine and lysine. *Journal of Biological Chemistry*. **279**, 45329-45336.

Vercammen, D., Belenghi, B., Van De, C. B., Beunens, T., Gavigan, J. A., De Rycke, R., Brackenier, A., Inze, D., Harris, J. L., Van Breusegem, F. (2006). Serpin1 of *Arabidopsis thaliana* is a suicide inhibitor for metacaspase 9. *Journal of Molecular Biology*. **364**, 625-636.

Vickerman, K., Luckins, A. G. (1969). Localization of variable antigens in the surface coat of *Trypanosoma brucei* using ferritin conjugated antibody. *Nature*. **224**, 1125-1126.

Vickerman, K. (1985). Developmental cycles and biology of pathogenic trypanosomes. *British Medical Bulletin*. **41**, 105-114.

Vincendeau, P., Bouteille, B. (2006). Immunology and immunopathology of African trypanosomiasis. *Anais da Academia Brasileira de Ciencia*s. **78**, 645-665.

Vodnik, M., Zager, U., Strukelj, B., Lunder, M. (2011). Phage display: selecting straws instead of a needle from a haystack. *Molecules*. **16**, 790-817.

Vreysen, M. J., Saleh, K. M., Ali, M. Y., Abdulla, A. M., Zhu, Z. R., Juma, K. G., Dyck, V. A., Msangi, A. R., Mkonyi, P. A., Feldmann, H. U. (2000). *Glossina austeni* (Diptera: *Glossinidae*) eradicated on the island of Unguja, Zanzibar, using the sterile insect technique. *Journal of economic entomology*. **93**, 123-135.

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**, S15-S25.

Walker, J. M. (1996). The bicinchoninic acid (BCA) assay for protein quantitation, In: The Protein Protocols Handbook, (Walker, J. M. (Ed.), Humana Press.

Walker, N. P., Talanian, R. V., Brady, K. D., Dang, L. C., Bump, N. J., Ferenz, C. R., Franklin, S., Ghayur, T., Hackett, M. C., Hammill, L. D. (1994). Crystal structure of the cysteine protease interleukin-1β-converting enzyme: A $(p20/p10)_2$ homodimer. *Cell*. **78**, 343-352.

Wamboga, C., Matovu, E., Bessell, P. R., Picado, A., Biéler, S., Ndung'u, J. M. (2017). Enhanced passive screening and diagnosis for *gambiense* human African trypanosomiasis in north-western Uganda - Moving towards elimination. *PLoS ONE*. **12**, e0186429.

Wang, L. F., Yu, M. (2009). Epitope mapping using phage-display random fragment libraries. *Methods in Molecular Biology*. **524**, 315-332.

Watanabe, N., Lam, E. (2005). Two *Arabidposis* metacaspases AtMCP1b and AtMCP2b are arginine/lysine-specific cysteine proteases and activate apoptosis-like cell death in yeast. *Journal of Biological Chemistry*. **280**, 14691-14699.

Watanabe, N., Lam, E. (2011). Calcium-dependent activation and autolysis of *Arabidopsis* metacaspase 2d. *Journal of Biological Chemistry*. **286**, 10027-10040.

Waugh, D. S. (2005). Making the most of affinity tags. *Trends in Biotechnology*. **23**, 316-320.

Webster, R. E. (1996). Biology of the filamentous bacteriophages, In: Phage display of peptides and proteins, a laboratory manual, (Kay, B. K., Winter, G., McCafferty, J. (Eds.)): United States, California, Academic Press.

Welburn, S. C., Févre, E. M., Coleman, P. G., Odiit, M., Maudlin, I. (2001a). Sleeping sickness: a tale of two diseases. *Trends in Parasitology*. **17**, 19-24.

Welburn, S. C., Picozzi, K., Fèvre, E. M., Coleman, P. G., Odiit, M., Maudlin, I. (2001b). Identification of human infective trypanosomes in animal reservoir of sleeping sickness in Uganda by means of serum-resistance-associated (SRA) gene. *Lancet*. **358**, 2017-2019.

Welburn, S. C., Coleman, P. G., Févre, E. M., Maudlin, I., Odiit, M., Eisler, M. C. (2006a). Crisis, what crisis? Control of Rhodesian sleeping sickness. *Trends in Parasitology*. **22**, 123-128.

Welburn, S. C., Macleod, E., Figgarella, K., Duszenko, M. (2006b). Progammed cell death in African trypanosomes. *Parasitology*. **132**, S7-S18.

Wemmer, S., Mashau, C., Fehrsen, J., van Wyngaardt, W., Du Plessis, D. H. (2010). Chicken scFvs and bivalent scFv-C_H fusions directed against HSP65 *Mycobacterium bovis*. *Biologicals*. **38**, 407-414.

Wen, S., Ma, Q. M., Zhang, Y. L., Yang, J. P., Zhao, G. H., Fu, D. Q., Luo, Y. B., Qu, G. Q. (2013). Biochemical evidence of key residues for the activation and autoprocessing of tomato type II metacaspase. *FEBS Letters*. **587**, 2517-2522.

Wen, X., Mousa, J. J., Bates, J. T., Lamb, R. A., Crowe, J. E. J., Jardetzky, T. S. (2017). Structural basis for antibody cross-neutralization of respiratory syncytial virus and human metapneumovirus. *Nature Microbiology*. **2**, 16272.

Wickstead, B., Ersfeld, K., Gull, K. (2004). The small chromosomes of *Trypanosoma brucei* involved in antigenic variation are constructed around repetitive palindromes. *Genome Research*. **14**, 1014-1024.

Wiederstein, M., Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*. **35**, W407-W110.

Winter, G., Milstein, C. (1991). Man-made antibodies. *Nature*. **349**, 293-299.

Wong, A. H., Yan, C., Shi, Y. (2012). Crystal structure of the yeast metacaspase Yca1. *Journal of Biological Chemistry*. **28**, 29251-29259.

Woo, P. T. (1970). The heamocrit centrifuge technique for the diagnosis of African trypanosomiasis. *Acta Tropica*. **27**, 384-386.

Wrzaczek, M., Vainonen, J. P., Stael, S., Tsiatsiani, L., Help-Rinta-Rahko, H., Gauthier, A., Kaufholdt, D., Bollhöner, B., Lamminmäki, A., Staes, A., Gevaert, K., Tuominen, H.,

Van Breusegem, F., Helariutta, Y., Kangasjärvi, J. (2015). GRIM REAPER peptide binds to receptor kinase PRK5 to trigger cell death in Arabidopsis. *EMBO Journal*. **34**, 55-66.

Wyatt, P. G., Gilbert, I. H., Read, K. D., Fairlamb, A. H. (2011). Target validation: linking target and chemical properties to desired product profile. *Current Topics in Medicinal Chemistry*. **11**, 1275-1283.

Yadav, S. C., Kumar, R., Manuja, A., Goyal, L., Gupta, A. K. (2012). Early detection of *Trypanosoma evansi* infection and monitoring of antibody levels by ELISA following treatment. *Journal of Parasitic Diseases*. **38**, 124-127.

Yang, J. M., Chen, C. C. (2004). GEMDOCK: a generic evolutionary method for molecular docking. *Proteins*. **55**, 288-304.

Yang, S. Y. (2010). Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discovery Today*. **15**, 444-450.

Yuriev, E., Agostino, M., Ramsland, P. A. (2009). Challenges and advances in computational docking: 2009 in review. *Journal of Molecular Recognition*. **24**, 149-164.

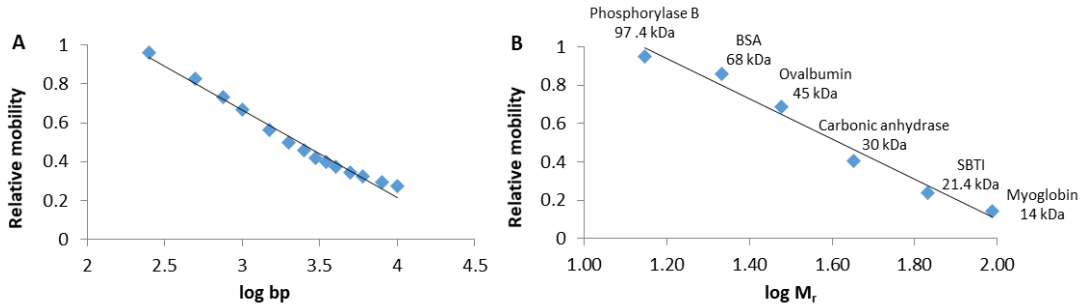
Zalila, H., González, I. J., El-Fadili, A. K., Delgado, M. B., Desponds, C., Schaff, C., Fasel, N. (2011). Processing of metacaspase into a cytoplasmic catalytic domain mediating cell death in *Leishmania major*. *Molecular Microbiology*. **79**, 222-239.

CONTENTS OF APPENDICES

APPENDICES A

Molecular weight marker calibrations

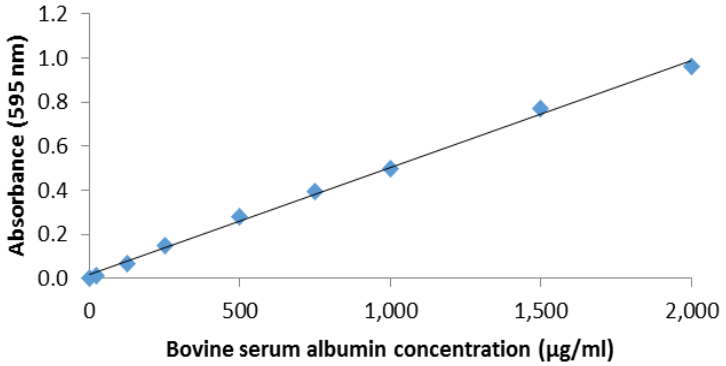
Standard curves were constructed for the molecular weight markers separated on a 1% (w/v) agarose gel (Appendix A1, panel A) and on a 12.5% reducing SDS-PAGE gel (panel B). These curves were constructed by the relation of the relative mobility of each of the marker proteins or DNA fragments to the log of its respective relative molecular mass (M_r) or number of base pairs (bp) respectively.



Appendix A1: Standard curves of the molecular weight markers used in agarose and SDS-PAGE electrophoresis. (A) The commercial O'GeneRuler™ 1 kb DNA ladder was used for 1% (w/v) agarose electrophoresis. The equation of the trendline is given by y = -0.449x + 2.102, with a correlation coefficient of 0.981. **(B)** The molecular weight marker used for reducing 12.5% SDS-PAGE electrophoresis consisted of phophorylase B (97.4 kDa), bovine serum albumin (BSA, 68 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), soyabean trypsin inhibitor (SBTI, 21.5 kDa) and lysozyme (14 kDa). The equation of the trendline is given by y= -1.050x + 2.198, with a correlation coefficient of 0.977.

Protein quantitation

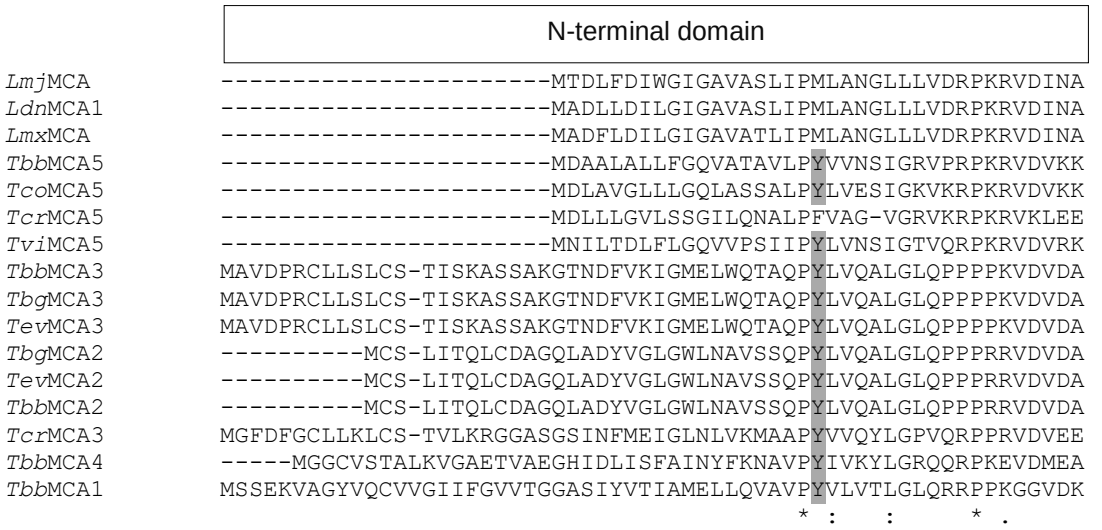
A water soluble, purple coloured complex is formed when cuprous ions are chelated by bicinchoninic acid (BCA). This complex is stable over a broad range of protein concentrations with a linear absorbance at 562 nm (Smith *et al.*, 1985). When compared to bovine serum albumin standards, the protein concentration in samples can be determined through the direct correlation between the protein concentration and the production of cuprous ions (Walker, 1996). Protein quantitation using the bicinchoninate reagent is superior over other methods due to its high tolerance towards resistance to detergents and denaturing reagents (Bradford, 1976), high sensitivity and low protein-to-protein variation and its stability in alkaline conditions which allows for a simplified one-step reaction (Walker, 1996). A BSA standard curve using the BCA™ protein assay kit is shown in Appendix A2.

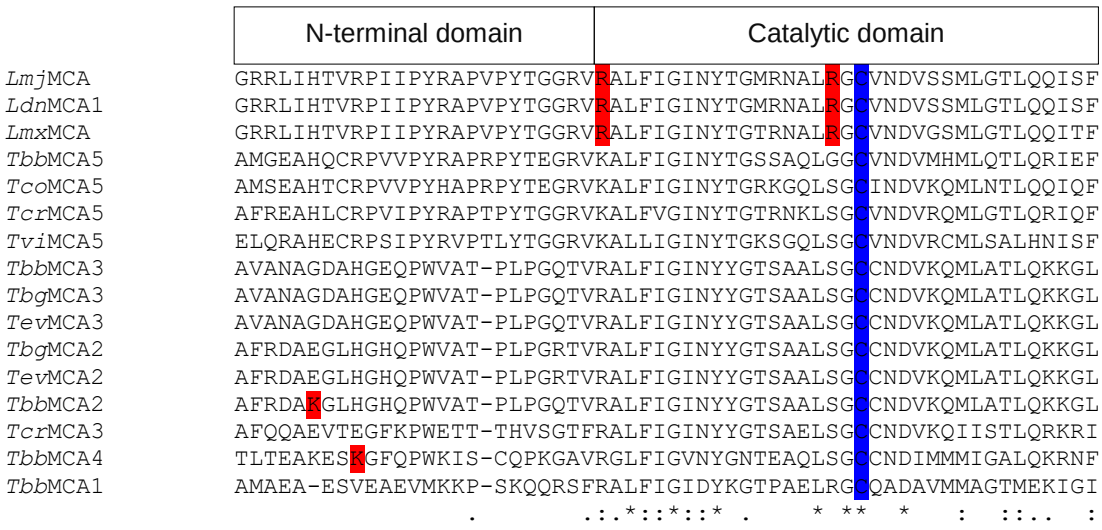


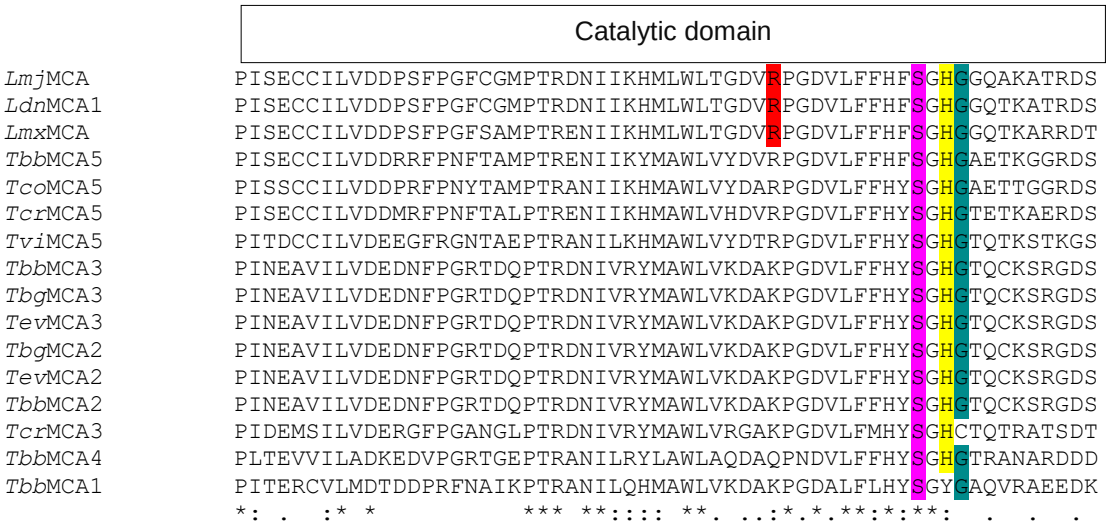
Appendix A2: Standard curve obtained using the BCA™ protein assay kit. Bovine serum albumin standards ranging from 25 to 2000 µg/ml were added to the BCA reagent and the resulting absorbance values measured at 595 nm after incubation at 37°C for 30 min. The equation of the trendline is given by y = 0.0005x + 0.016 with a correlation coefficient of 0.997.

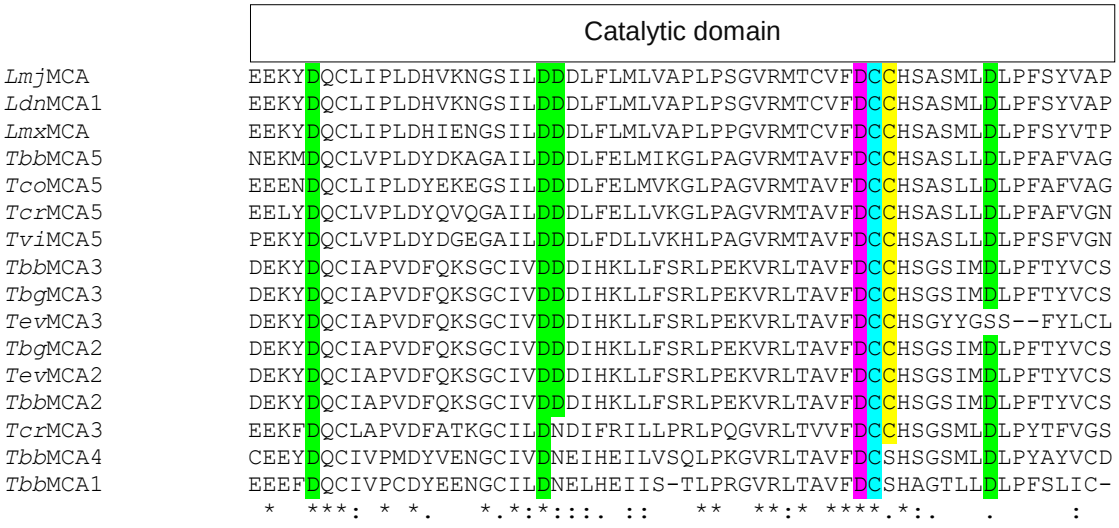
Relatedness and conservation of the catalytic residues of the kinetoplastid MCAs

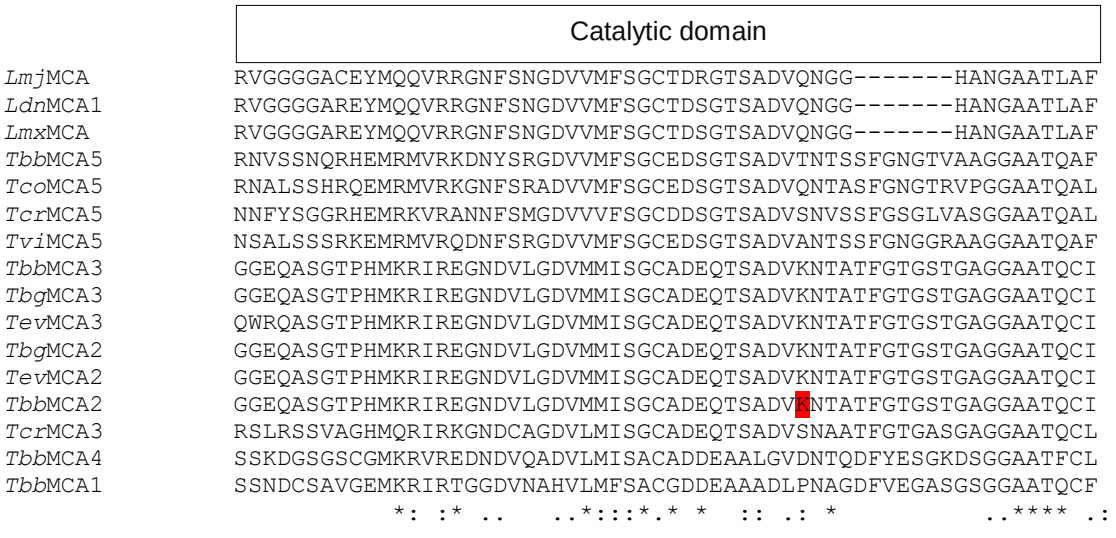
The MCAs -1, -4 and -5 of *T. b. brucei, T. b. gambiense* and *T. evansi* share a 100% identity (Fig. 2.1). As such, only the analyses of the MCAs-1, -4 and -5 of *T. b. brucei* will be reported. The *Lmj*MCA and *Ldn*MCA2 share a 99% identity and as such, only *Lmj*MCA will be reported.

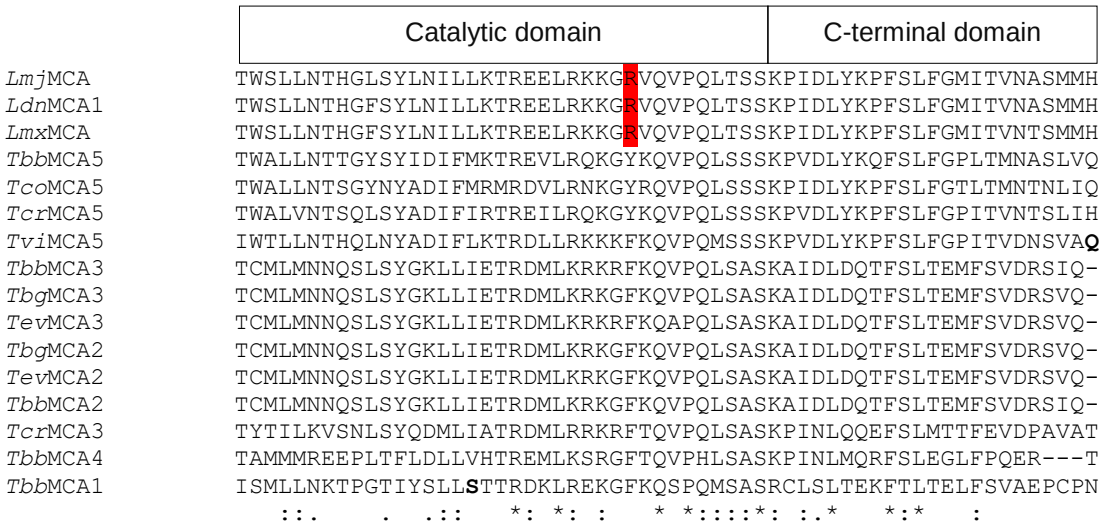


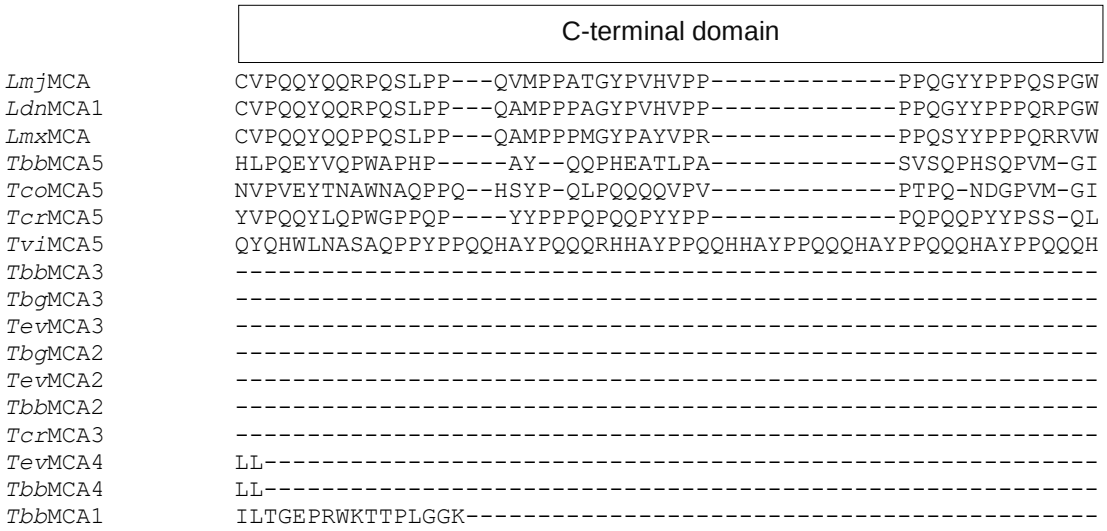


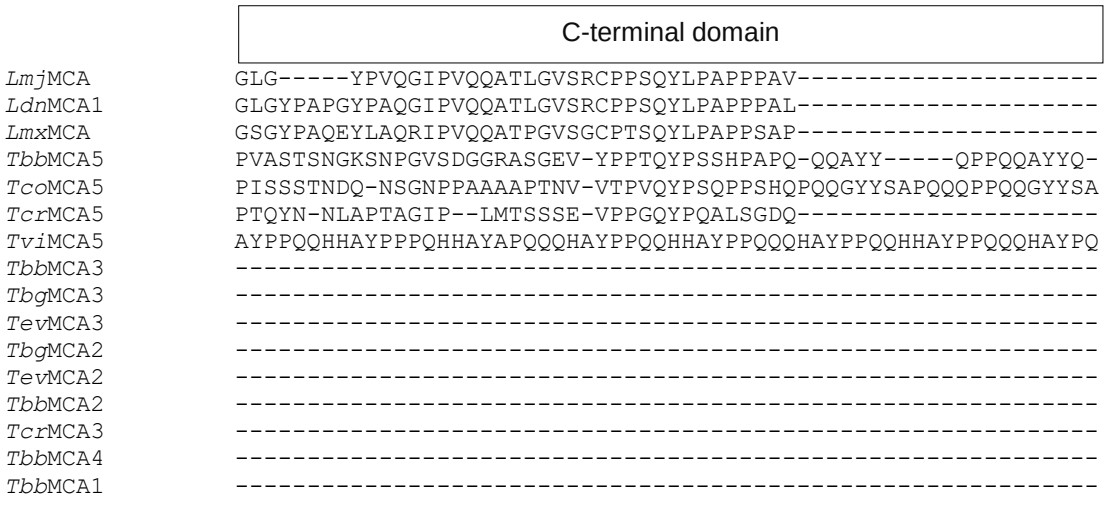

Catalytic domain
LmjMCA PISECCILVDDPSFPGFCGMPTRDNIIKHMLWLTGDVRPGDVLFFHFSGHGGQAKATRDS
LdnMCA1 PISECCILVDDPSFPGFCGMPTRDNIIKHMLWLTGDVRPGDVLFFHFSGHGGQTKATRDS
LmxMCA PISECCILVDDPSFPGFSAMPTRENIIKHMLWLTGDVRPGDVLFFHFSGHGGQTKARRDT
TbbMCA5 PISECCILVDDRRFPNFTAMPTRENIIKYMAWLVYDVRPGDVLFFHFSGHGAETKGGRDS
TcoMCA5 PISSCCILVDDPRFPNYTAMPTRANIIKHMAWLVYDARPGDVLFFHYSGHGAETTGGRDS
TcrMCA5 PISECCILVDDMRFPNFTALPTRENIIKHMAWLVHDVRPGDVLFFHYSGHGTETKAERDS
TviMCA5 PITDCCILVDEEGFRGNTAEPTRANILKHMAWLVYDTRPGDVLFFHYSGHGTQTKSTKGS
TbbMCA3 PINEAVILVDEDNFPGRTDQPTRDNIVRYMAWLVKDAKPGDVLFFHYSGHGTQCKSRGDS
TbgMCA3 PINEAVILVDEDNFPGRTDQPTRDNIVRYMAWLVKDAKPGDVLFFHYSGHGTQCKSRGDS
TevMCA3 PINEAVILVDEDNFPGRTDQPTRDNIVRYMAWLVKDAKPGDVLFFHYSGHGTQCKSRGDS
TbgMCA2 PINEAVILVDEDNFPGRTDQPTRDNIVRYMAWLVKDAKPGDVLFFHYSGHGTQCKSRGDS
TevMCA2 PINEAVILVDEDNFPGRTDQPTRDNIVRYMAWLVKDAKPGDVLFFHYSGHGTQCKSRGDS
TbbMCA2 PINEAVILVDEDNFPGRTDQPTRDNIVRYMAWLVKDAKPGDVLFFHYSGHGTQCKSRGDS
TcrMCA3 PIDEMSILVDERGFPGANGLPTRDNIVRYMAWLVRGAKPGDVLFMHYSGHCTQTRATSDT
TbbMCA4 PLTEVVILADKEDVPGRTGEPTRANILRYLAWLAQDAQPNDVLFFHYSGHGTRANARDDD
TbbMCA1 PITERCVLMDTDDPRFNAIKPTRANILQHMAWLVKDAKPGDALFLHYSGYGAQVRAEEDK

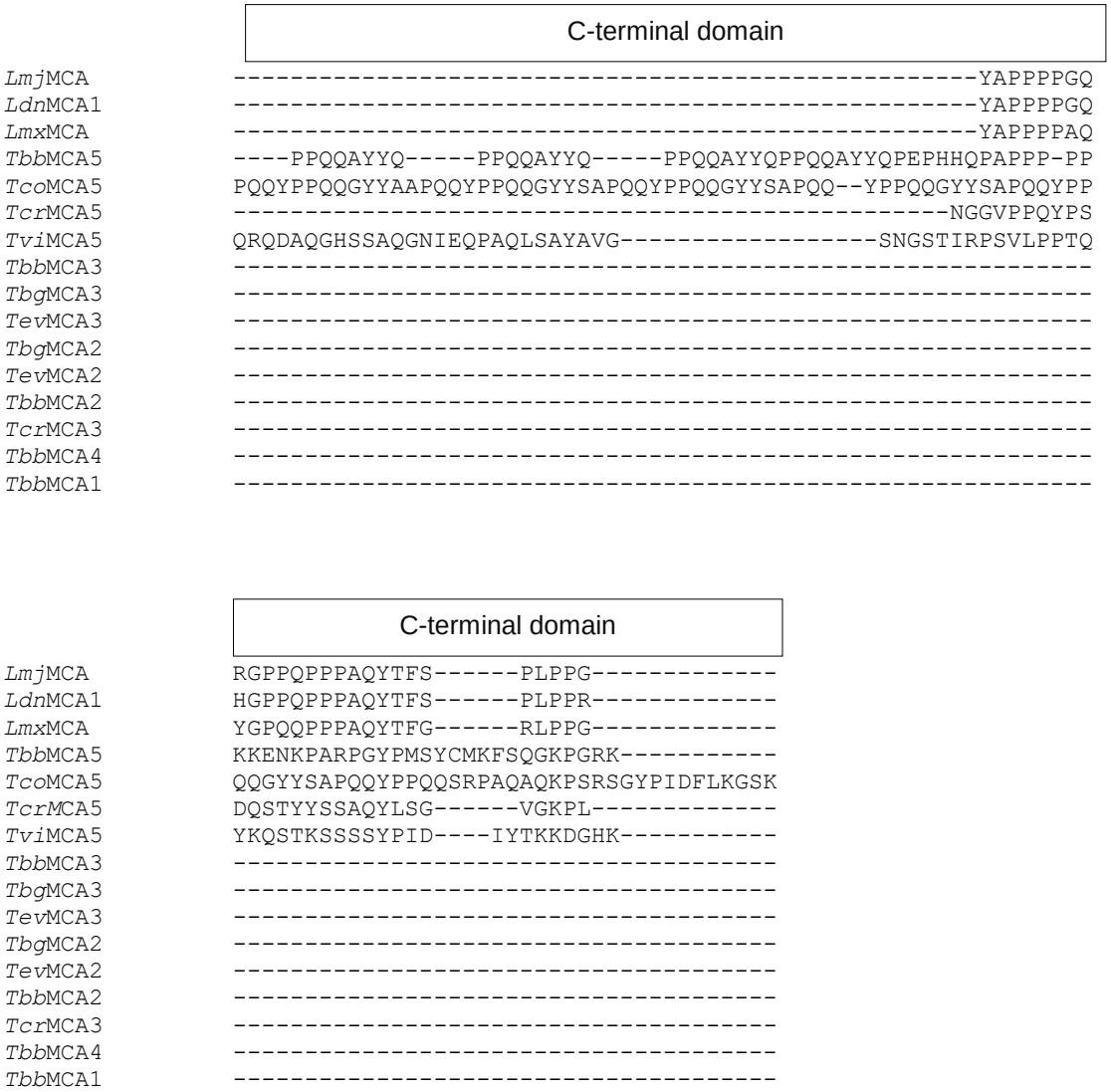

Catalytic domain
LmjMCA EEKYDQCLIPLDHVKNGSILDDDLFLMLVAPLPSGVRMTCVFDCCHSASMLDLPFSYVAP
LdnMCA1 EEKYDQCLIPLDHVKNGSILDDDLFLMLVAPLPSGVRMTCVFDCCHSASMLDLPFSYVAP
LmxMCA EEKYDQCLIPLDHIENGSILDDDLFLMLVAPLPPGVRMTCVFDCCHSASMLDLPFSYVTP
TbbMCA5 NEKMDQCLVPLDYDKAGAILDDDLFELMIKGLPAGVRMTAVFDCCHSASLLDLPFAFVAG
TcoMCA5 EEENDQCLIPLDYEKEGSILDDDLFELMVKGLPAGVRMTAVFDCCHSASLLDLPFAFVAG
TcrMCA5 EELYDQCLVPLDYQVQGAILDDDLFELLVKGLPAGVRMTAVFDCCHSASLLDLPFAFVGN
TviMCA5 PEKYDQCLVPLDYDGEGAILDDDLFDLLVKHLPAGVRMTAVFDCCHSASLLDLPFSFVGN
TbbMCA3 DEKYDQCIAPVDFQKSGCIVDDDIHKLLFSRLPEKVRLTAVFDCCHSGSIMDLPFTYVCS
TbgMCA3 DEKYDQCIAPVDFQKSGCIVDDDIHKLLFSRLPEKVRLTAVFDCCHSGSIMDLPFTYVCS
TevMCA3 DEKYDQCIAPVDFQKSGCIVDDDIHKLLFSRLPEKVRLTAVFDCCHSGYYGSS--FYLCL
TbgMCA2 DEKYDQCIAPVDFQKSGCIVDDDIHKLLFSRLPEKVRLTAVFDCCHSGSIMDLPFTYVCS
TevMCA2 DEKYDQCIAPVDFQKSGCIVDDDIHKLLFSRLPEKVRLTAVFDCCHSGSIMDLPFTYVCS
TbbMCA2 DEKYDQCIAPVDFQKSGCIVDDDIHKLLFSRLPEKVRLTAVFDCCHSGSIMDLPFTYVCS
TcrMCA3 EEKFDQCLAPVDFATKGCILDNDIFRILLPRLPQGVRLTVVFDCCHSGSMLDLPYTFVGS
TbbMCA4 CEEYDQCIVPMDYVENGCIVDNEIHEILVSQLPKGVRLTAVFDCSHSGSMLDLPYAYVCD
TbbMCA1 EEEFDQCIVPCDYEENGCILDNELHEIIS-TLPRGVRLTAVFDCSHAGTLLDLPFSLIC-


Catalytic domain
LmjMCA RVGGGGACEYMQQVRRGNFSNGDVVMFSGCTDRGTSADVQNGG-------HANGAATLAF
LdnMCA1 RVGGGGAREYMQQVRRGNFSNGDVVMFSGCTDSGTSADVQNGG-------HANGAATLAF
LmxMCA RVGGGGAREYMQQVRRGNFSNGDVVMFSGCTDSGTSADVQNGG-------HANGAATLAF
TbbMCA5 RNVSSNQRHEMRMVRKDNYSRGDVVMFSGCEDSGTSADVTNTSSFGNGTVAAGGAATQAF
TcoMCA5 RNALSSHRQEMRMVRKGNFSRADVVMFSGCEDSGTSADVQNTASFGNGTRVPGGAATQAL
TcrMCA5 NNFYSGGRHEMRKVRANNFSMGDVVVFSGCDDSGTSADVSNVSSFGSGLVASGGAATQAL
TviMCA5 NSALSSSRKEMRMVRQDNFSRGDVVMFSGCEDSGTSADVANTSSFGNGGRAAGGAATQAF
TbbMCA3 GGEQASGTPHMKRIREGNDVLGDVMMISGCADEQTSADVKNTATFGTGSTGAGGAATQCI
TbgMCA3 GGEQASGTPHMKRIREGNDVLGDVMMISGCADEQTSADVKNTATFGTGSTGAGGAATQCI
TevMCA3 QWRQASGTPHMKRIREGNDVLGDVMMISGCADEQTSADVKNTATFGTGSTGAGGAATQCI
TbgMCA2 GGEQASGTPHMKRIREGNDVLGDVMMISGCADEQTSADVKNTATFGTGSTGAGGAATQCI
TevMCA2 GGEQASGTPHMKRIREGNDVLGDVMMISGCADEQTSADVKNTATFGTGSTGAGGAATQCI
TbbMCA2 GGEQASGTPHMKRIREGNDVLGDVMMISGCADEQTSADVKNTATFGTGSTGAGGAATQCI
TcrMCA3 RSLRSSVAGHMQRIRKGNDCAGDVLMISGCADEQTSADVSNAATFGTGASGAGGAATQCL
TbbMCA4 SSKDGSGSCGMKRVREDNDVQADVLMISACADDEAALGVDNTQDFYESGKDSGGAATFCL
TbbMCA1 SSNDCSAVGEMKRIRTGGDVNAHVLMFSACGDDEAAADLPNAGDFVEGASGSGGAATQCF


Catalytic domain
C-terminal domain
LmjMCA TWSLLNTHGLSYLNILLKTREELRKKGRVQVPQLTSSKPIDLYKPFSLFGMITVNASMMH
LdnMCA1 TWSLLNTHGFSYLNILLKTREELRKKGRVQVPQLTSSKPIDLYKPFSLFGMITVNASMMH
LmxMCA TWSLLNTHGFSYLNILLKTREELRKKGRVQVPQLTSSKPIDLYKPFSLFGMITVNTSMMH
TbbMCA5 TWALLNTTGYSYIDIFMKTREVLRQKGYKQVPQLSSSKPVDLYKQFSLFGPLTMNASLVQ
TcoMCA5 TWALLNTSGYNYADIFMRMRDVLRNKGYRQVPQLSSSKPIDLYKPFSLFGTLTMNTNLIQ
TcrMCA5 TWALVNTSQLSYADIFIRTREILRQKGYKQVPQLSSSKPVDLYKPFSLFGPITVNTSLIH
TviMCA5 IWTLLNTHQLNYADIFLKTRDLLRKKKFKQVPQMSSSKPVDLYKPFSLFGPITVDNSVAQ
TbbMCA3 TCMLMNNQSLSYGKLLIETRDMLKRKRFKQVPQLSASKAIDLDQTFSLTEMFSVDRSIQ-
TbgMCA3 TCMLMNNQSLSYGKLLIETRDMLKRKGFKQVPQLSASKAIDLDQTFSLTEMFSVDRSVQ-
TevMCA3 TCMLMNNQSLSYGKLLIETRDMLKRKRFKQAPQLSASKAIDLDQTFSLTEMFSVDRSVQ-
TbgMCA2 TCMLMNNQSLSYGKLLIETRDMLKRKGFKQVPQLSASKAIDLDQTFSLTEMFSVDRSVQ-
TevMCA2 TCMLMNNQSLSYGKLLIETRDMLKRKGFKQVPQLSASKAIDLDQTFSLTEMFSVDRSVQ-
TbbMCA2 TCMLMNNQSLSYGKLLIETRDMLKRKGFKQVPQLSASKAIDLDQTFSLTEMFSVDRSIQ-
TcrMCA3 TYTILKVSNLSYQDMLIATRDMLRRKRFTQVPQLSASKPINLQQEFSLMTTFEVDPAVAT
TbbMCA4 TAMMMREEPLTFLDLLVHTREMLKSRGFTQVPHLSASKPINLMQRFSLEGLFPQER---T
TbbMCA1 ISMLLNKTPGTIYSLLSTTRDKLREKGFKQSPQMSASRCLSLTEKFTLTELFSVAEPCPN


C-terminal domain
LmjMCA CVPQQYQQRPQSLPP---QVMPPATGYPVHVPP----------PPQGYYPPPQSPGW
LdnMCA1 CVPQQYQQRPQSLPP---QAMPPPAGYPVHVPP----------PPQGYYPPPQRPGW
LmxMCA CVPQQYQQPPQSLPP---QAMPPPMGYPAYVPR----------PPQSYYPPPQRRVW
TbbMCA5 HLPQEYVQPWAPHP-----AY--QQPHEATLPA----------SVSQPHSQPVM-GI
TcoMCA5 NVPVEYTNAWNAQPPQ--HSYP-QLPQQQQVPV----------PTPQ-NDGPVM-GI
TcrMCA5 YVPQQYLQPWGPPQP----YYPPPQPQQPYYPP----------PQPQQPYYPSS-QL
TviMCA5 QYQHWLNASAQPPYPPQQHAYPQQQRHHAYPPQQHHAYPPQQQHAYPPQQQHAYPPQQQH
TbbMCA3 ----------
TbgMCA3 ----------
TevMCA3 ----------
TbgMCA2 ----------
TevMCA2 ----------
TbbMCA2 ----------
TcrMCA3 ----------
TevMCA4 LL----------
TbbMCA4 LL----------
TbbMCA1 ILTGEPRWKTTPLGGK----------

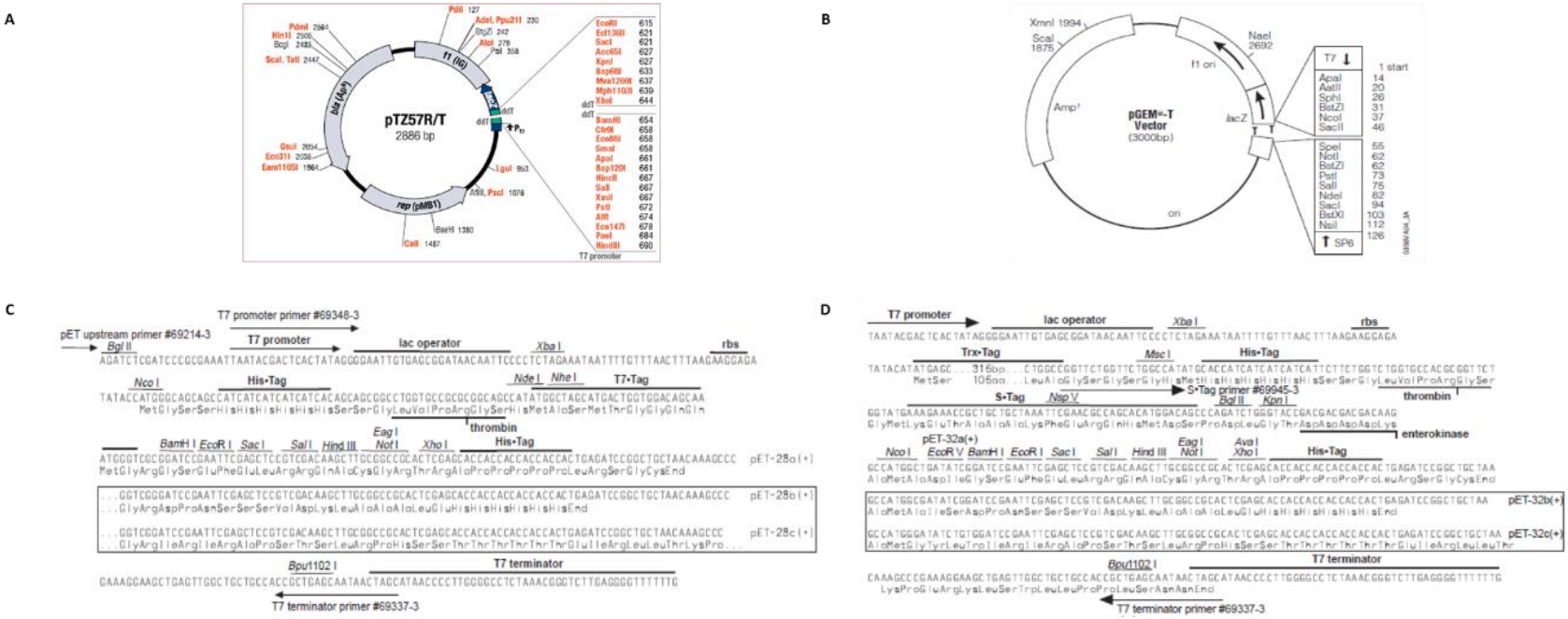

C-terminal domain
LmjMCA GLG-----YPVQGIPVQQATLGVSRCPPSQYLPAPPPAV----------
LdnMCA1 GLGYPAPGYPAQGIPVQQATLGVSRCPPSQYLPAPPPAL----------
LmxMCA GSGYPAQEYLAQRIPVQQATPGVSGCPTSQYLPAPPSAP----------
TbbMCA5 PVASTSNGKSNPGVSDGGRASGEV-YPPTQYPSSHPAPQ-QQAYY-----QPPQQAYYQ-
TcoMCA5 PISSSTNDQ-NSGNPPAAAAPTNV-VTPVQYPSQPPSHQPQQGYYSAPQQQPPQQGYYSA
TcrMCA5 PTQYN-NLAPTAGIP--LMTSSSE-VPPGQYPQALSGDQ----------
TviMCA5 AYPPQQHHAYPPPQHHAYAPQQQHAYPPQQHHAYPPQQQHAYPPQQHHAYPPQQQHAYPQ
TbbMCA3 ----------
TbgMCA3 ----------
TevMCA3 ----------
TbgMCA2 ----------
TevMCA2 ----------
TbbMCA2 ----------
TcrMCA3 ----------
TbbMCA4 ----------
TbbMCA1 ----------



Appendix A3: Multiple sequence alignment of the kinetoplastid MCAs. The protein sequences were obtained from the TriTrypDB (Aslett *et al.*, 2010). *Tbb*MCA from *T. b. brucei* (927), *Tbg*MCA from *T. b. gambiense* (DAL972), *Tvi*MCA from *T. vivax* (Y486), *Tev*MCA from *T. evansi* (STIB 805), *Tco*MCA from *T. congolense* (IL3000), *Tcr*MCA from *T. cruzi* (Sylvio X10/1), *Lmj*MCA from *L. major* (Friedlin), *Lmx*MCA from *L. mexicana* (MHOM/GT/2001/U1103), *Ldn*MCA from *L. donovani* (BPK282A1) and aligned using ClustalW (Larkin *et al.*, 2007). Alignment characters are annotated as follows: conserved residues (*), strongly similar properties (:) and weakly similar properties (.). The amino acid residues involved in the S_1 binding pocket are highlighted in pink, those in the catalytic dyad in yellow, adjacent Cys to the catalytic Cys in turquoise, the secondary catalytic Cys in blue and the calcium binding sites in green. The gatekeeper residue is highlighted in grey and the sites of autoprocessing, reported in literature, in red. The residue predicted to be involved in the formation of the oxyanion hole is highlighted in teal.

Cloning and expression plasmids used for *TbbMCA2* and *TviMCA5*

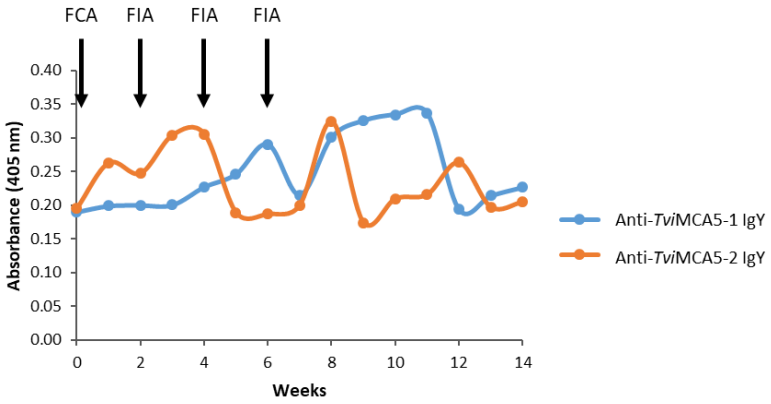
The plasmid maps of the cloning (panels A and B) and expression (panels C and D) plasmids used in the present study is given in Appendix A4.



Appendix A4: Sequence maps of the multiple cloning sites of the plasmids used of the cloning of *Tbb*MCA2 and *Tvi*MCA5. T/A cloning plasmids **(A)** pTZ57R/t and **(B)** pGem®-T along with the **(C)** pET-28a and **(D)** pET-32a expression plasmids which both code for a N-terminal 6xHis tagged recombinant protein.

Chicken anti-*Tvi*MCA5 IgY production

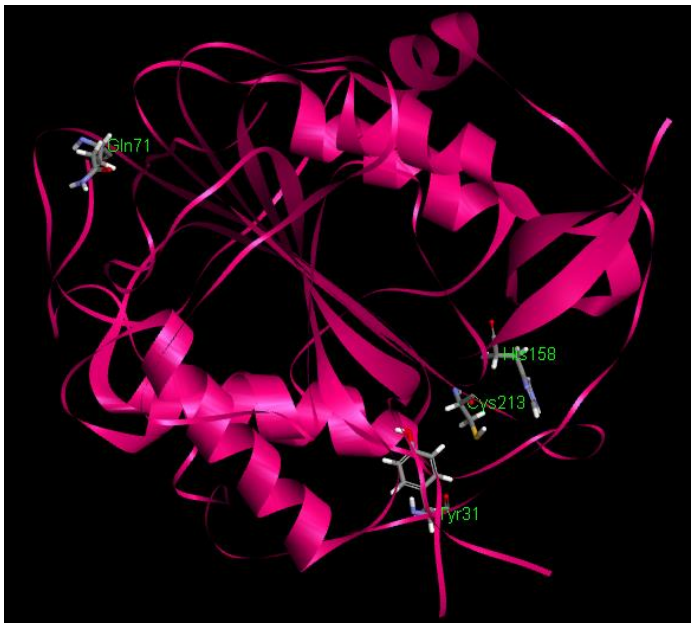
Chickens were immunised with *Tvi*MCA5 triturated with Freund's complete adjuvant (FCA) at week 0 and thereafter with Freund's incomplete adjuvant (FIA) at weeks 2, 4 and 6. The pre-immune IgY is denoted by week 0 which had a low absorbance value. Antibody production fluctuated throughout the immunisation period (Appendix A5). In chicken 1, antibody production was above that of the pre-immune at weeks 4 to 6, 8 to 11, and for chicken 2, weeks 1 to 4, 8 and 13.



Appendix A5: ELISA of anti-*Tvi*MCA5 IgY antibodies isolated from the egg yolks of immunised chickens. ELISA plates were coated with *Tvi*MCA5 (1 µg/ml in PBS, pH 7.2), blocked with 0.5% (w/v) BSA-PBS, and incubated with anti-*Tvi*MCA5 IgY from chickens 1 and 2, weeks 1 to 14 (100 µg/ml). Rabbit anti-chicken IgY HRPO conjugate (1:5 000) and ABTS·H_2O_2 were used as the detection system. The absorbance readings at 405 nm represent the average of duplicate experiments at 60 min development.

Mutation in the *TbbMCA2* amplified gene

Using the 3D structure of *Tbb*MCA2 (PDB ID: 4AF8), solved by X-ray diffraction, the catalytic dyad residues, His158 and Cys213, which had been mutated to a Gly, were identified along with the gatekeeper reside Tyr31 which operates as a switch to allow access into the catalytic pocket (McLuskey *et al.*, 2012). After sequencing the high-fidelity taq amplified *TbbMCA2* gene from the DNA of *T. b. brucei* (927), that the Gln71 residue was found to be amplified as Arg71. This mutation was modelled on the protein structure to determine where in relation it was to the active site and if it could affect substrate binding (Appendix A6).



Appendix A6: The 3D structure of *Tbb*MCA2 4AF8 showing the mutation which was introduced after PCR in relation to the catalytic dyad. The 3D structure of *Tbb*MCA2^{C213G} (4AF8), from diffraction data collected at 1.4 Å (McLuskey *et al.*, 2012), was modelled using Discovery Studio software (Dassault Systèmes BIOVIA), and the mutated Gln71 residue along with the catalytic dyad, His158 and Gly213, and gate keeper (Tyr31) residues were shown in a stick configuration.

APPENDICES B

Primers designed for mutagenesis

The primers for the generation of the catalytic dyad mutants, H147A and C202G, along with the adjacent Cys, C201G, were designed using the Agilent QuikChange primer design software and detailed in Appendix B1.

Appendix B1: Primer sequences to be used in the mutagenesis studies of *Tco*MCA5 and *Tvi*MCA5.

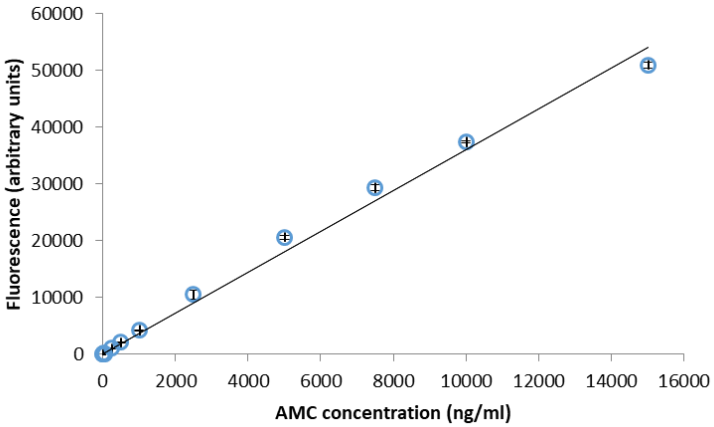
Primer	5' - 3'
TcoMCA5^{H147A} forward	TTC ACT ACT CTG GG**G CC**G GTG CGG AGA CGA C
TcoMCA5^{H147A} reverse	GTC GTC TCC GCA CC**G GC**C CCA GAG TAG TGA A
TcoMCA5^{C201G} forward	ATG ACG GCC GTC TTT GAT **GGC** TGC CAC TCT GCC
TcoMCA5^{C201G} reverse	GGC AGA GTG GCA **GCC** ATC AAA GAC GGC CGT CAT
TcoMCA5^{C202G} forward	GGC CGT CTT TGA TTG C**GG C**CA CTC TGC CTC TCT TC
TcoMCA5^{C202G} reverse	GAA GAG AGG CAG AGT G**GC C**GC AAT CAA AGA CGG CC
TviMCA5^{H147A} forward	CGT GTT GTT TTT TCA TTA TTC TGG A**GC C**GG CAC TCA AAC TAA AAG TAC
TviMCA5^{H147A} reverse	GTA CTT TTA GTT TGA GTG CC**G GC**T CCA GAA TAA TGA AAA AAC AAC ACG
TviMCA5^{C201G} forward	GTG CGT ATG ACA GCT GTT TTT GAT **GGC** TGT CAT TCA GCC TCT CTT TTG G
TviMCA5^{C201G} reverse	CCA AAA GAG AGG CTG AAT GAC A**GG C**AT CAA AAA CAG CTG TCA TAC GCA C
TviMCA5^{C202G} forward	GCG TAT GAC AGC TGT TTT TGA TTG T**GG C**CA TTC AGC CTC TCT TTT GG
TviMCA5^{C202G} reverse	CCA AAA GAG AGG CTG AAT G**GC C**AC AAT CAA AAA CAG CTG TCA TAC GC

Sequences in bold correspond to the point mutations.

AMC standard curve

To quantify the hydrolysis of fluorogenic peptide substrates by the purified recombinant and native MCAs, an AMC calibration curve was constructed. The AMC standards ranging from 5 to 15 000 nM (50 µl) were incubated with MCA assay buffer (50 mM Tris-HCl buffer, pH 7.2, 150 mM NaCl, 10 mM $CaCl_2$, 5 mM DTT, 50 µl) at 37°C for 10 min and the fluorescence (Ex_{360nm} and Em_{460nm}) measured using a FLUORStar Optima Spectrophotometer from BMG Labtech (Offenburg, Germany) (Appendix B2).

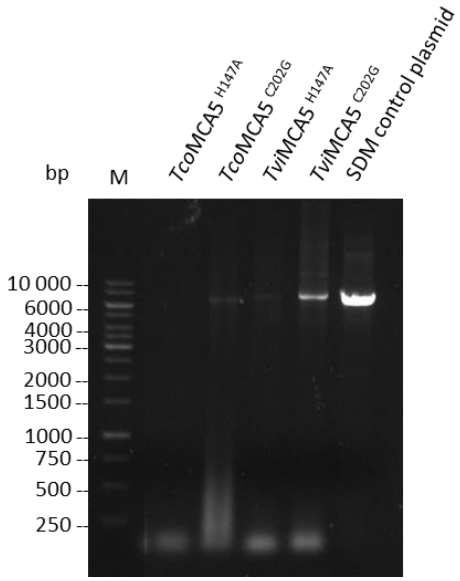
The slope of the calibration curve was used to determine the velocity of the enzymes when hydrolysing different substrates.



Appendix B2: AMC standard curve to link the fluorescence from the hydrolysis of fluorogenic peptide substrates by the MCAs to their velocity. Dilutions of AMC, in triplicate, were incubated with MCA assay buffer at for 10 min at 37°C and the excitation and emission fluorescence measured at 360 and 460 nm respectively. The equation of the trendline is given by y = 3.6124x with a correlation coefficient of 0.9909.

Analysis of the generated megaprimer after mutagenesis

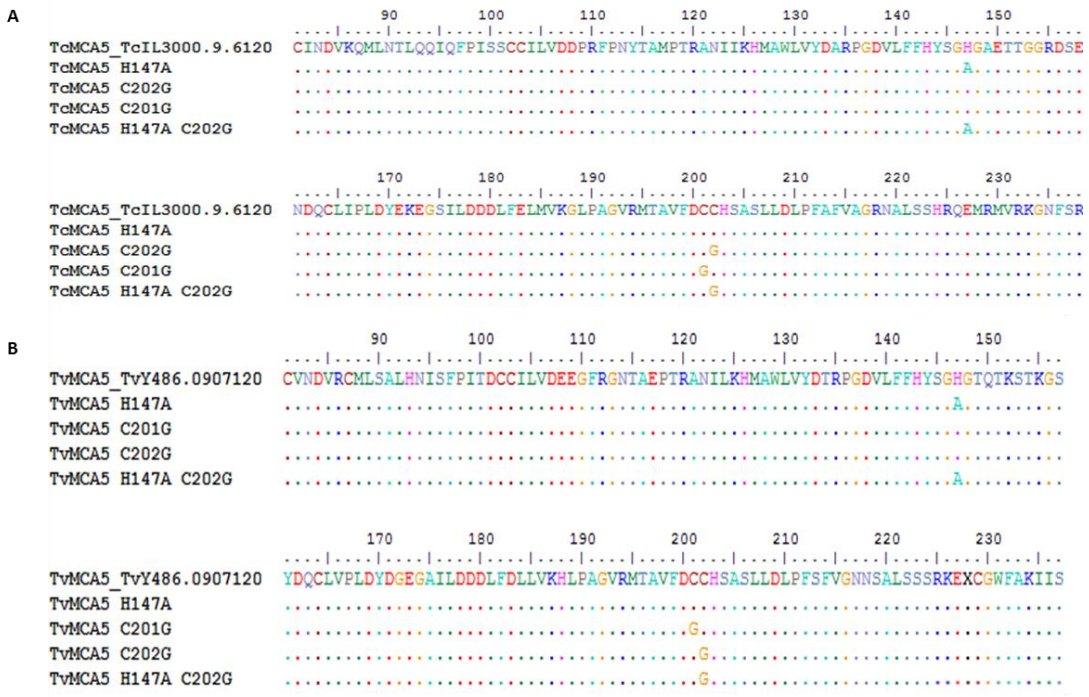
Using the primers shown in Appendix B1, along with the plasmid DNA templates of the WT *Tco*MCA5-pET28a and *Tvi*MCA5-pET32a, the point mutation within the plasmid was performed using the Q5 high-fidelity master mix. Successful mutagenesis resulted in bands at approximately 7 000 and 7 500 bp for *Tco*MCA5 and *Tvi*MCA5 respectively. Only a few mutants are shown in Appendix B3.



Appendix B3: High-fidelity amplification of *TcoMCA5* and *TviMCA5* mutants using the megaprimer method. Samples of the PCR amplification of point mutants were electrophoresed on a 1% (w/v) agarose gel containing 0.5 µg/ml ethidium bromide.

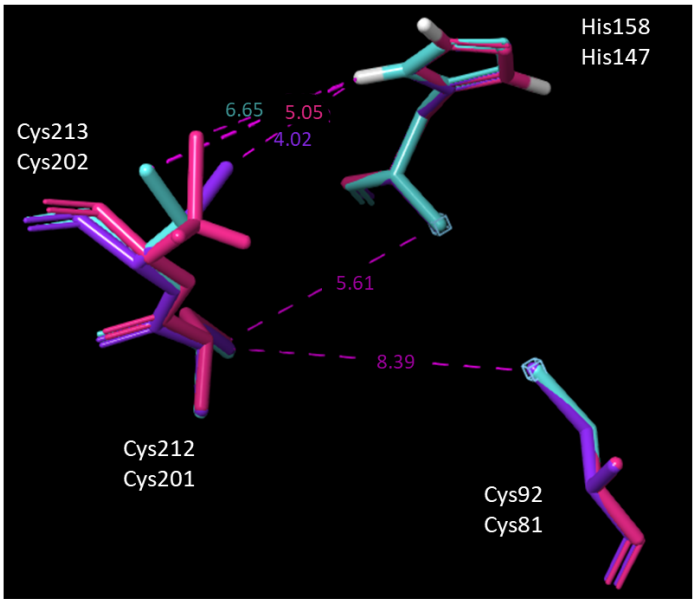
Alignment of sequenced WT and mutant *Tco*MCA5 and *Tvi*MCA5

To determine whether the resulting clones after transformation of the mutant plasmids, into competent NEB 5-alpha *E. coli* cells, possessed the desired mutations, the plasmid DNA was isolated and sequenced using the gene forward primer (Table 2.1). The results show that each of the selected clones were positive for the desired mutations (Appendix B4).



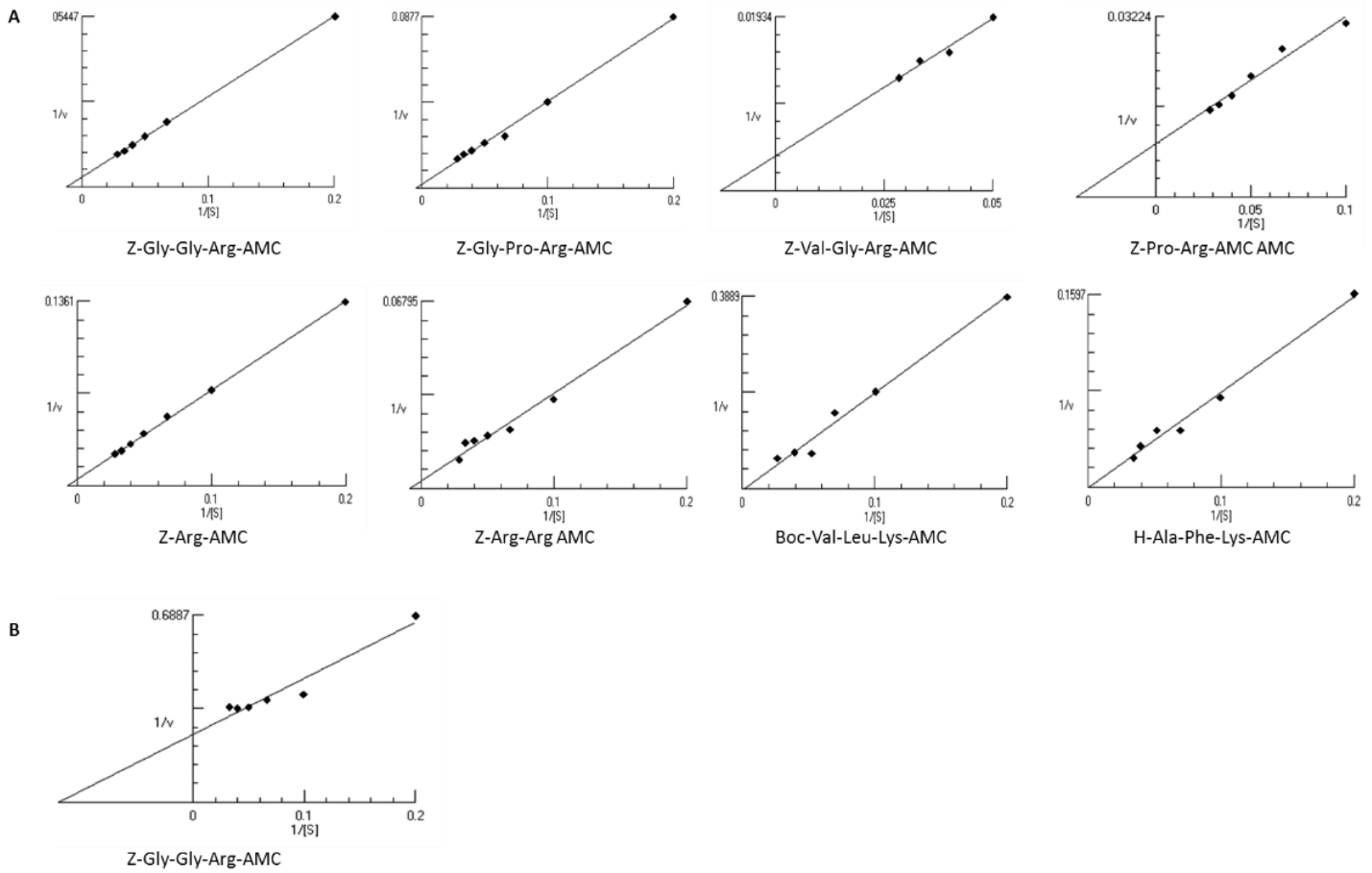
Appendix B4: Sequence alignment of the mutant clones of *Tco*MCA5 and *Tvi*MCA5. The plasmid DNA of **(A)** *Tco*MCA5 and **(B)** *Tvi*MCA5 mutants was sequenced using the forward gene primer at the Central Analytical Facilities (CAF), Stellenbosch University, and the data analysed using the BioEdit software (Hall, 1999) and compared to that from GenBank®.

Secondary catalytic Cys of *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5



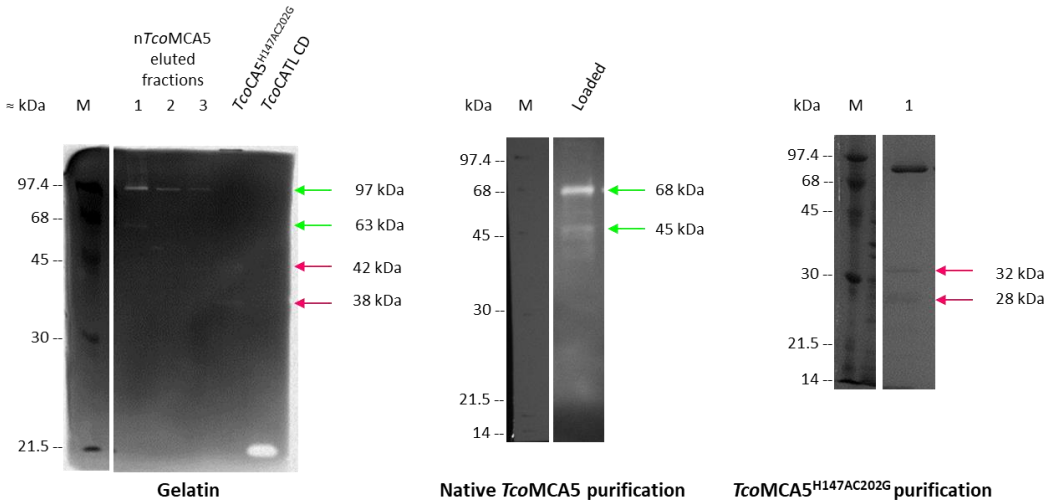
Appendix B5: Visualisation of the proximity of the catalytically important residues in *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5. The residues comprising the catalytic dyad (His158 and Cys213; His147 and Cys202), Cys adjacent to the catalytic Cys (Cys212 and Cys201) and the secondary catalytic Cys (Cys92 and Cys81) of *Tbb*MCA2 4AF8 (pink), *Tco*MCA5 (purple) and *Tvi*MCA5 (turquoise) were visualised using the Discovery Studio (Dassault Systèmes BIOVIA).

Lineweaver-Burke plots



Appendix B6: Lineweaver-Burke plots of the hydrolysis of various fluorogenic peptide substrates by recombinant *Tco*MCA5^{H147AC202G} and native *Tco*MCA5. Recombinant **(A)** *Tco*MCA5^{H147AC202G} (0.5 µg/ml) and **(B)** native *Tco*MCA5 (7.5 µg/ml) were incubated in MCA assay buffer before the addition of 0 to 30µM of various fluorogenic peptide substrates in triplicate and duplicate respectively. The resultant fluorescence from the hydrolysis of the AMC fluorophore, over 60 min, was measured at Ex_{360nm} and Em_{460nm}. After the outliers were removed, the kinetic parameters, K_m and V_{max} were determined using the Hyper32® software.

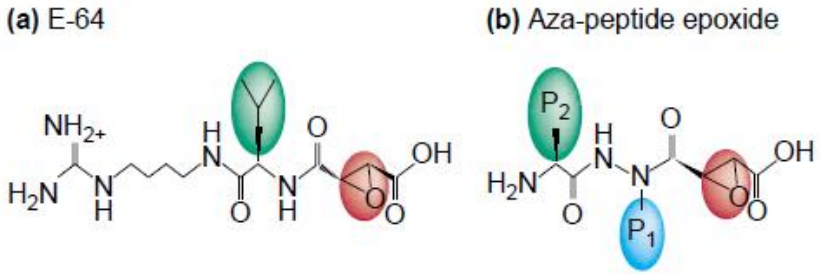
Molecular weight comparison between non-reducing zymograms and reducing SDS-PAGE gels



Appendix B7: Molecular weight comparison of active MCA5 from a non-reducing zymogram with MCA5 on a reducing SDS-PAGE gel. (A) A direct comparison of gelatin containing non-reducing SDS-PAGE gel, also known as a zymogram, to reducing SDS-PAGE gels of **(B)** immunoaffinity purified native *Tco*MCA5 (green) detected in a western blot, and **(C)** nickel affinity purified recombinant *Tco*MCA5^{H147AC202G} (pink).

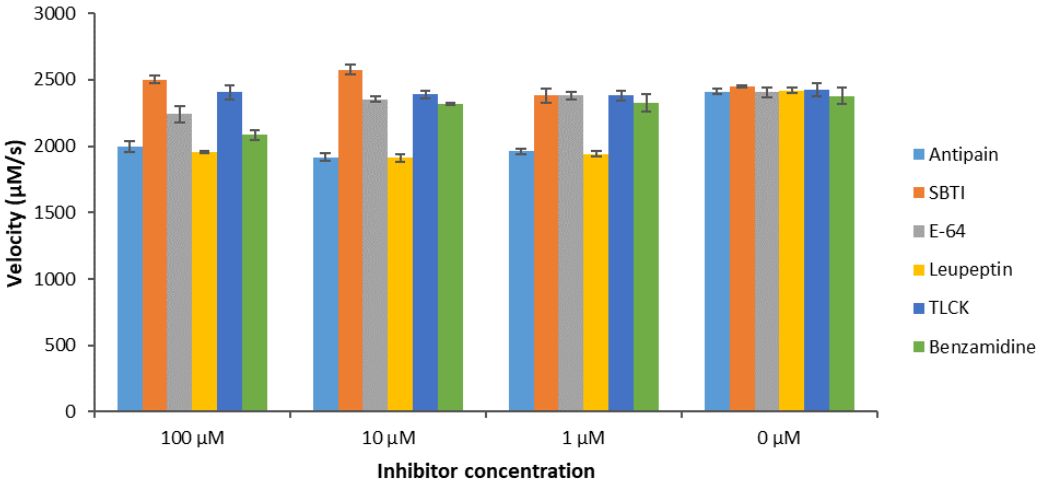
Clan CD peptidases resistance to classical E-64 inhibitor

Appendix B8 is taken directly from Mottram *et al.* (2003).



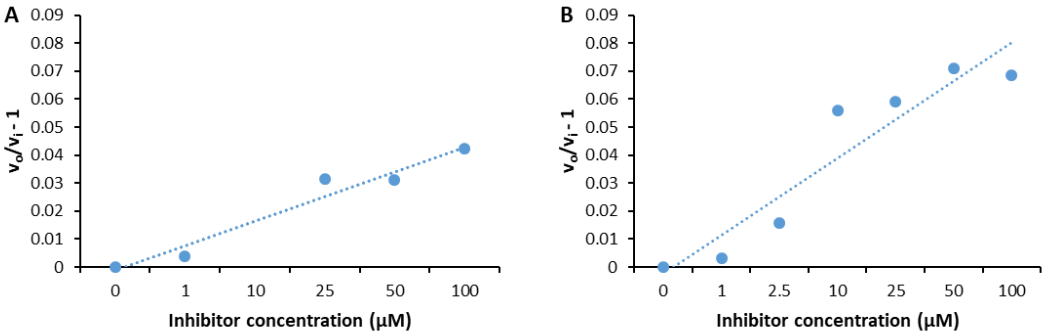
Appendix B8: Chemical structure of (a) the generic clan CA inhibitor E-64 and (b) aza-peptide epoxide inhibitors. The P_2 binding sites are indicated in green, which is a leucine residue in the case of E-64. The P_1 site of the aza-peptide epoxide inhibitors is indicated in blue. E-64 lacks a corresponding P_1 residue, which accounts for the compound's lack of activity towards clan CD peptidases. The active epoxide functional group, that covalently binds to the thiol of the active site Cys in cysteine peptidases, is shown in red.

Resultant velocity after inhibitor screening



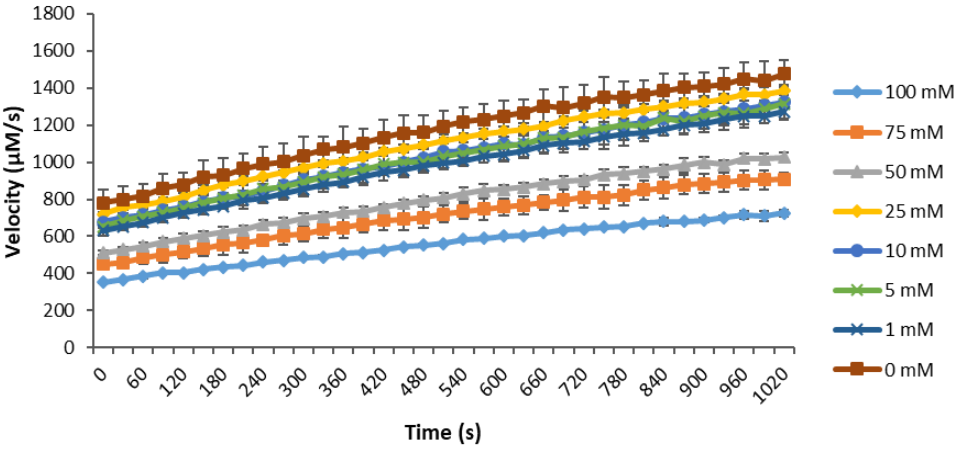
Appendix B9: Impact of varying concentrations of various inhibitors on the activity of recombinant *Tco*MCA5^{H147AC202G}. Recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) was incubated in MCA assay buffer containing 10 µM Z-Gly-Gly-Arg-AMC, before the addition of 1, 10 and 100 µM of various peptidase inhibitors, in triplicate, diluted in assay buffer without substrate. The resultant fluorescence from the hydrolysis of the AMC fluorophore, over 60 min, was measured at Ex_{360nm} and Em_{460nm}. No change in velocity was observed for calpain I, bestatin, AEBSF, 1,10 phenanthroline, Z-Phe-Ala-FMK, Z-Val-Ala-Asp-FMK, iodoacetic acid and iodoacetate.

Determination of inhibition kinetic constant



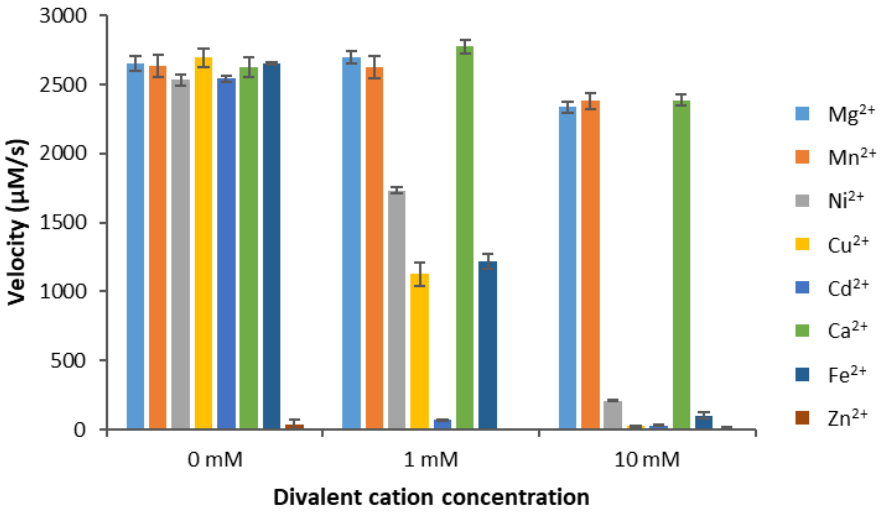
Appendix B10: Determination of the inhibition constants for antipain and leupeptin against recombinant *Tco*MCA5^{H147AC202G}. Using equation [D] in Section 3.2.6.5, the inhibition constant, K_i, for **(A)** antipain and **(B)** leupeptin was determined. The equation of the trendlines is given by **(A)** y = 0.0088x – 0. 01 with a correlation coefficient of 0.9538 and **(B)** y = 0.0137x -0.0159 with a correlation coefficient of 0.8910.

Effect of calcium



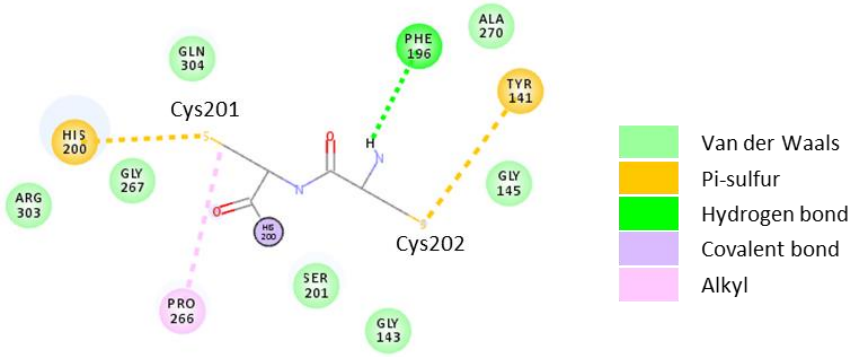
Appendix B11: The effect of varying concentrations of calcium on the velocity of recombinant *Tco*MCA5^{H147AC202G}. Recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) was incubated in MCA assay buffer containing 10 µM Z-Gly-Gly-Arg-AMC, before the addition of 0 to 100 mM $CaCl_2$, in triplicate, diluted in assay buffer without substrate. The resultant fluorescence from the hydrolysis of the AMC fluorophore, over 60 min, was measured at Ex_{360nm} and Em_{460nm}.

Effect of various divalent cations



Appendix B12: The effect of varying concentrations of various divalent cations on the velocity of recombinant *Tco*MCA5^{H147AC202G}. Recombinant *Tco*MCA5^{H147AC202G} (0.5 µg/ml) was incubated in MCA assay buffer containing 10 µM Z-Gly-Gly-Arg-AMC, before the addition of 1 and 10 mM of various divalent cations diluted in assay buffer without substrate. This was done in triplicate. The resultant fluorescence from the hydrolysis of the AMC fluorophore, over 60 min, was measured at Ex_{360nm} and Em_{460nm}.

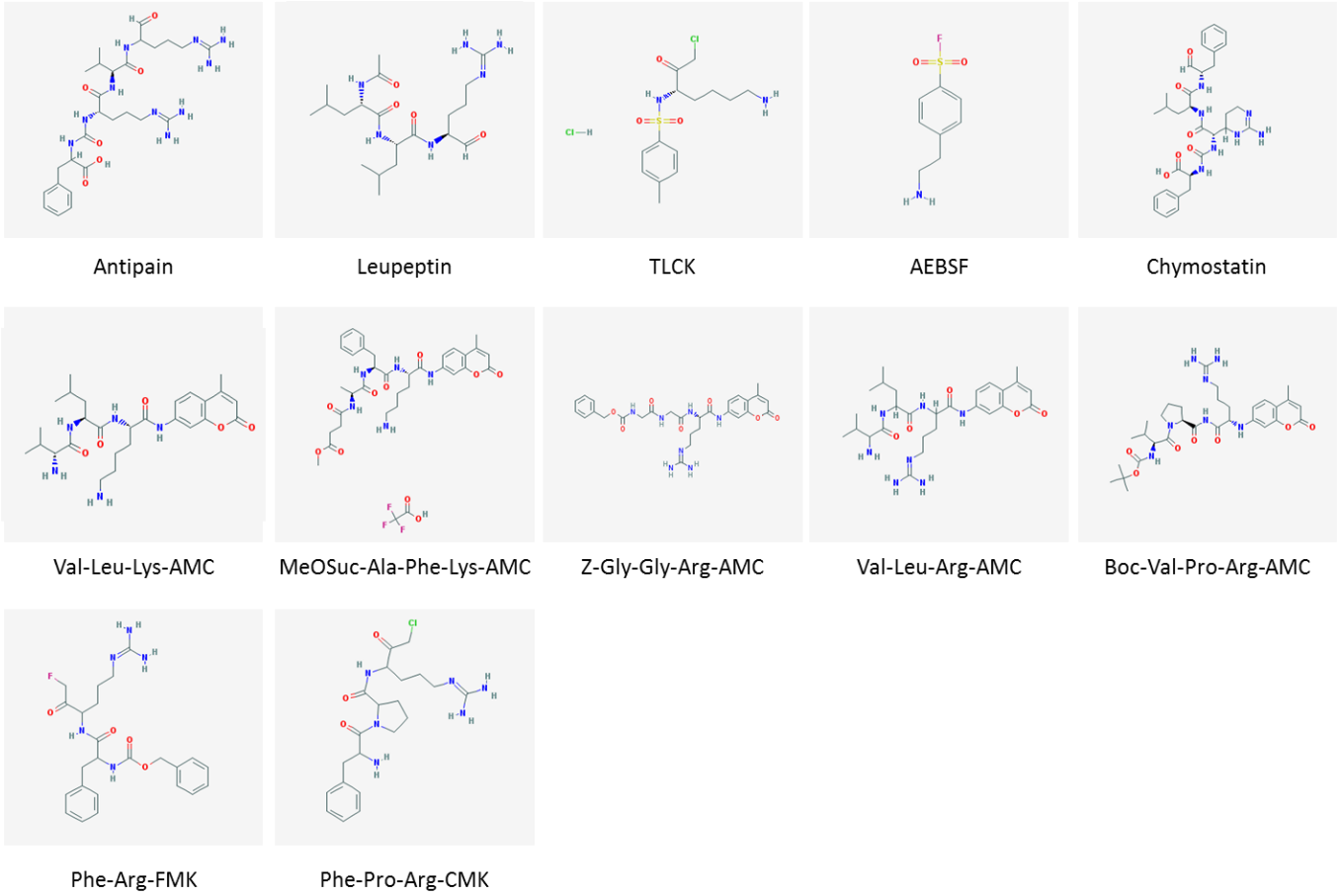
Interactions of the catalytic and adjacent Cys residues



Appendix B13: The interactions of the catalytic Cys202 and adjacent Cys201 in *Tco*MCA5. The 2D interaction map highlighting the Pi-sulfur interactions of the catalytic Cys with Tyr141 and Cys adjacent to the catalytic Cys, Cys201, with His200. The 2D interaction map for *Tvi*MCA5 is identical to that of *Tco*MCA5 and is thus omitted.

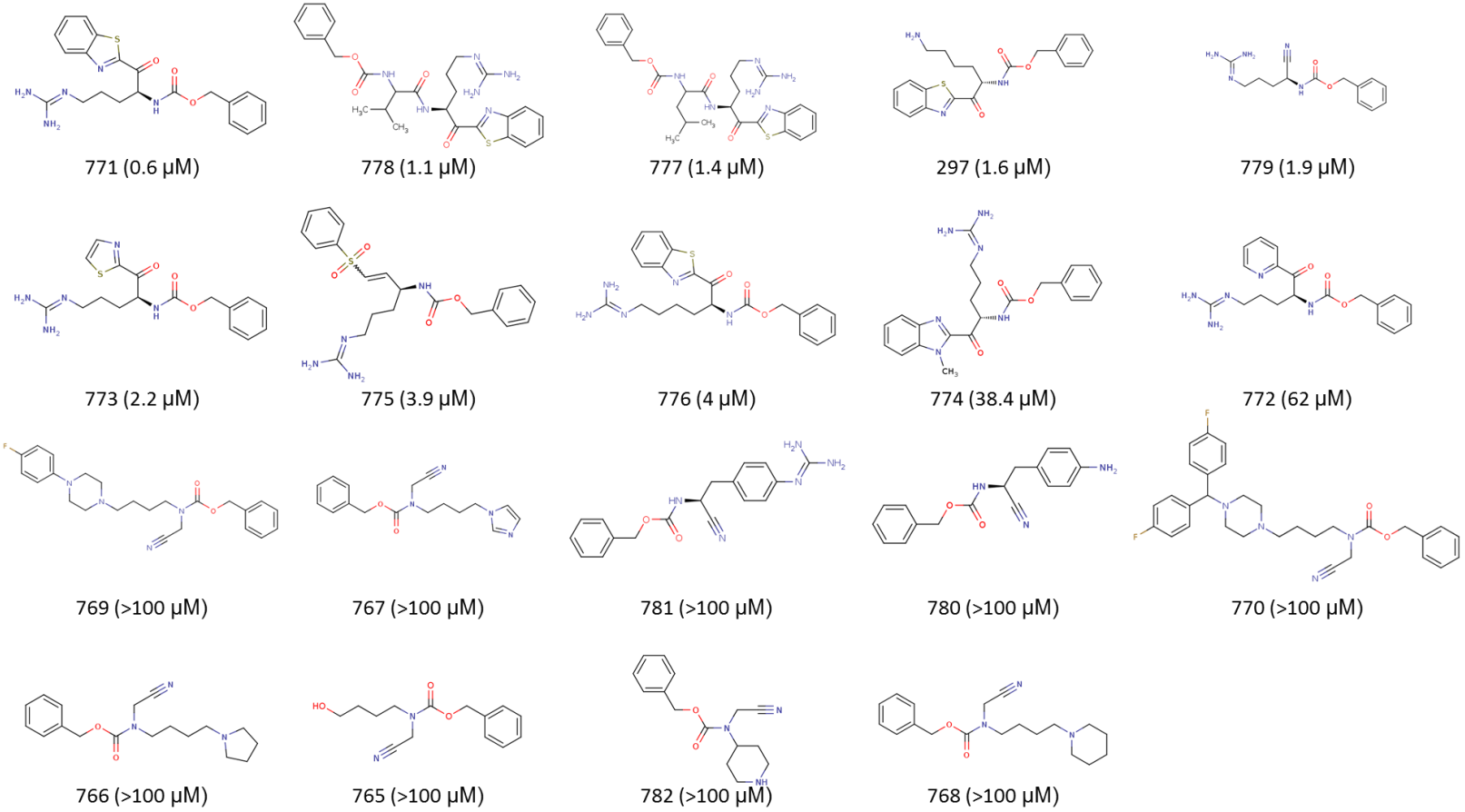
APPENDICES C

Commercial library



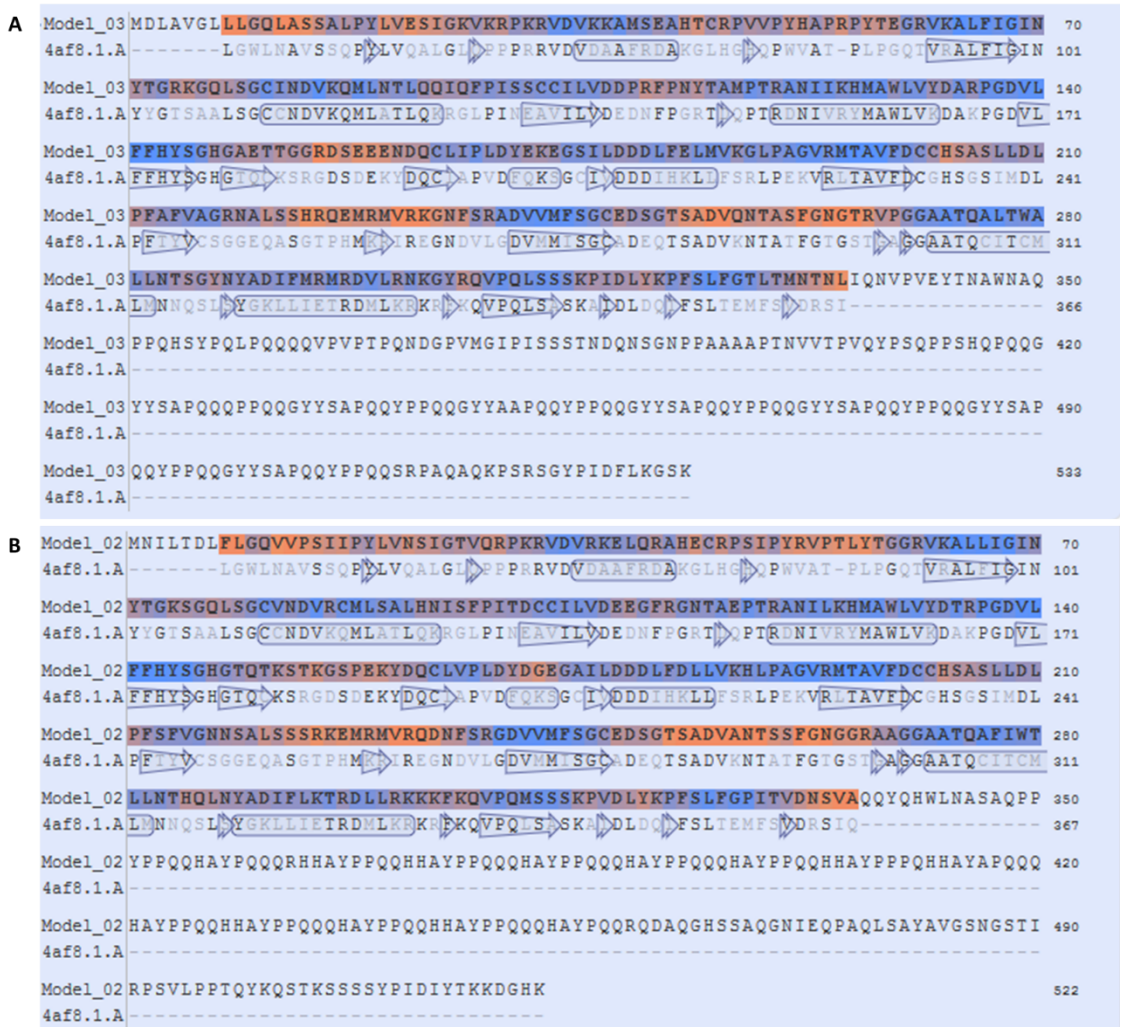
Appendix C1: Inhibitors and substrates used in literature for the characterisation of various MCAs. The chemical structures were obtained from the PubChem compound database (https://pubchem.ncbi.nlm.nih.gov/)(Kim *et al.*, 2016).

Berg library



Appendix C2: Inhibitors designed based on the P1 site preference of *Tbb*MCA2 for Arg and Lys residues. The chemical structures of the Berg library (Berg *et al.*, 2012) were obtained from the Binding Database (Gilson *et al.*, 2016)(http://www.bindingdb.org). The given numbers are the last three numbers from the bindingDB monomer ID. The IC_{50} values against recombinant *Tbb*MCA2 are indicated in brackets as per Berg *et al*. (2012).

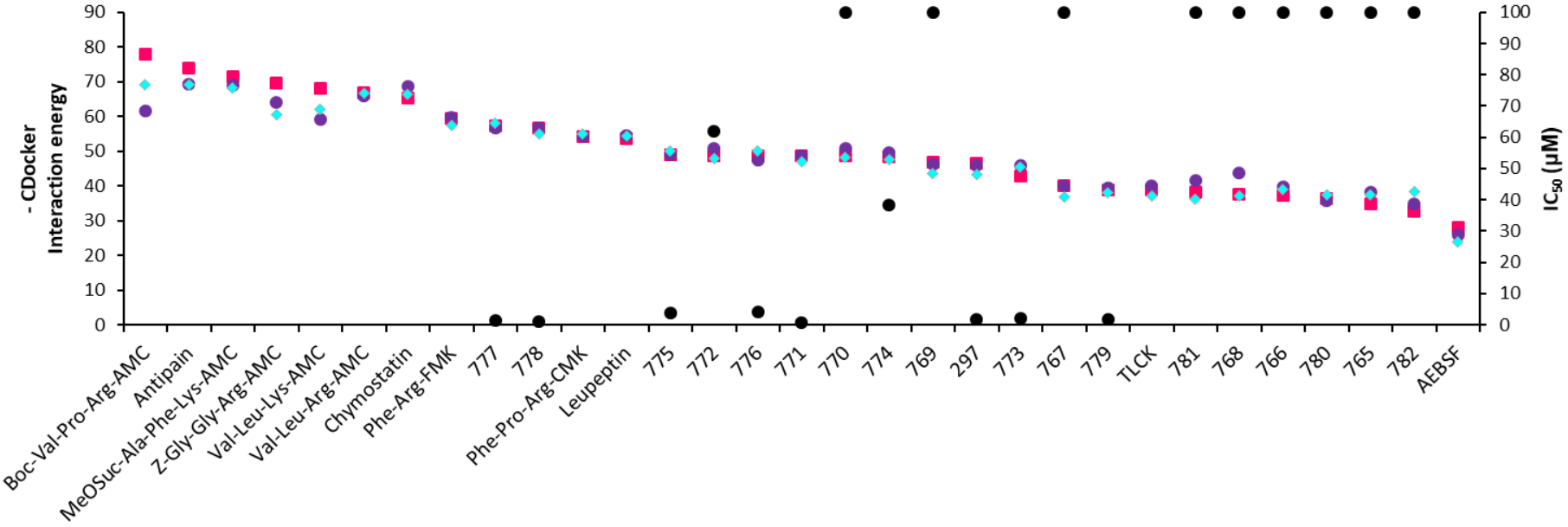
Enlarged sequence alignment of the *Tbb*MCA2 4AF8 template used for the homology modelling of *Tco*MCA5 and *Tvi*MCA5



Appendix C3: Homology modelling of *Tco*MCA5 and *Tvi*MCA5. The alignment of the **(A)** *TcoM*CA5 and **(B)** *Tvi*MCA5 sequences to the *Tbb*MCA2 4AF8 template is shown with the Qmean scores indicated in blue (high quality) and red/orange (low quality).

Correlation of the -CDocker interaction energies and the IC_{50} values of the Berg ligands

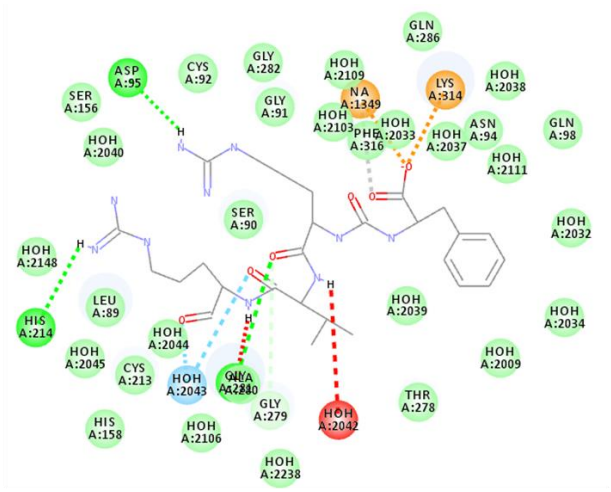
A total of 31 compounds, consisting of seven substrates, four inhibitors and 19 Berg ligands were docked into the active site of *Tbb*MCA2 4AF8, *Tco*MCA5 and *Tvi*MCA5, of which only the Berg ligands had an associated IC_{50} value. Of the Berg ligands, only ten possessed an IC_{50} value below 100 µM.



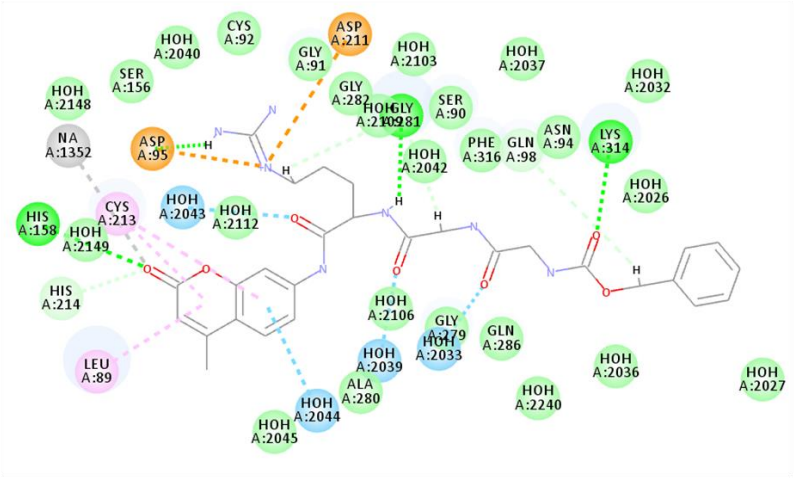
Appendix C4: Correlation of experimental IC_{50} values of the Berg ligands and the -CDocker interaction energies of ligands docked into the active site of *Tbb*MCA2, *Tco*MCA5 and *Tvi*MCA5. The alignment of the *Tbb*MCA2 (pink), *TcoM*CA5 (purple) and *Tvi*MCA5 (turquoise) and the IC_{50} values (black) as determined for the Berg ligands against *Tbb*MCA2 (Berg *et al.*, 2012).

2D interactions of ligands docked in the active site of *Tbb*MCA2 4AF8

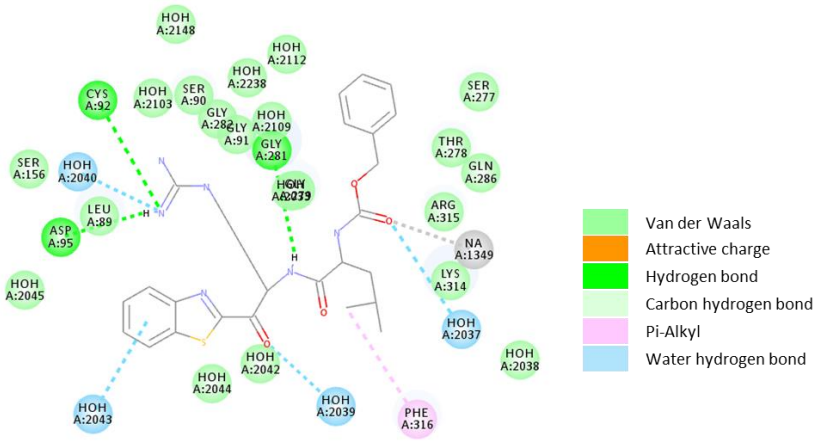
Antipain



Z-Gly-Gly-Arg-AMC

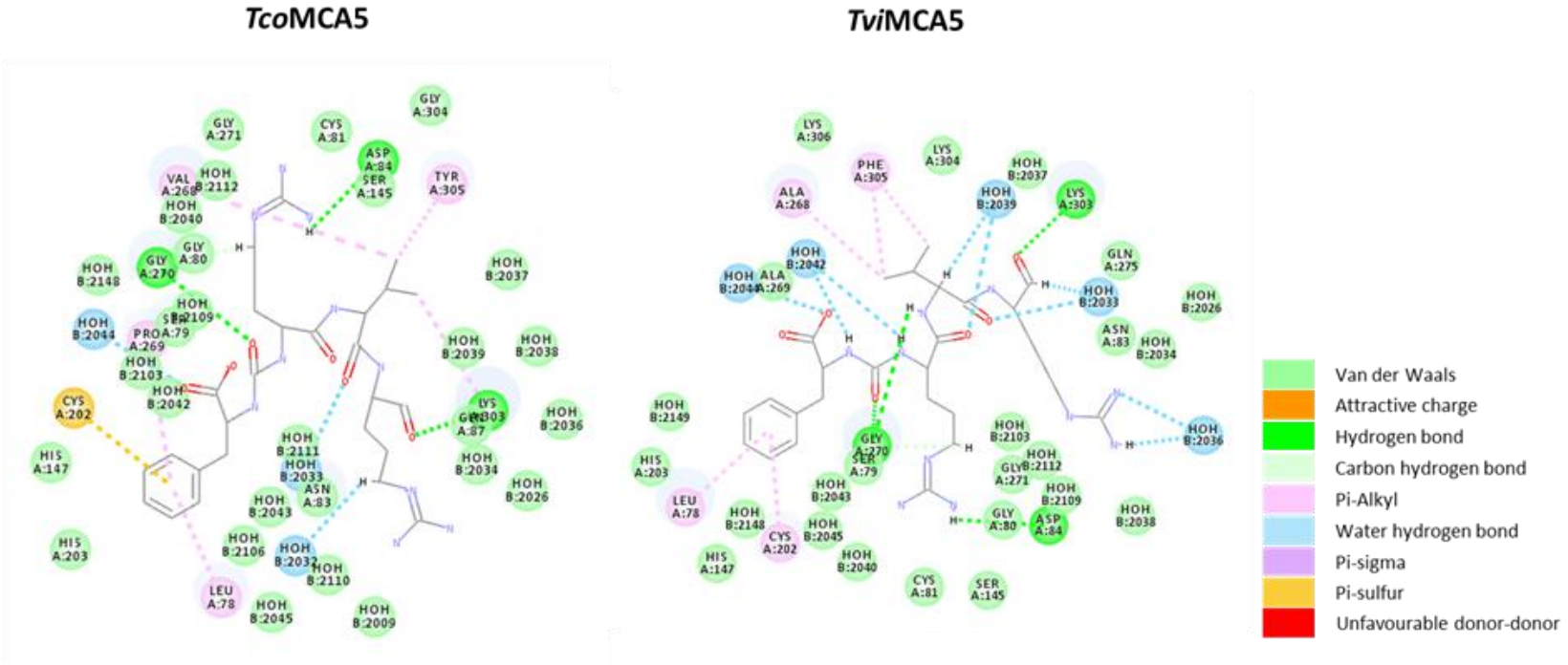


777

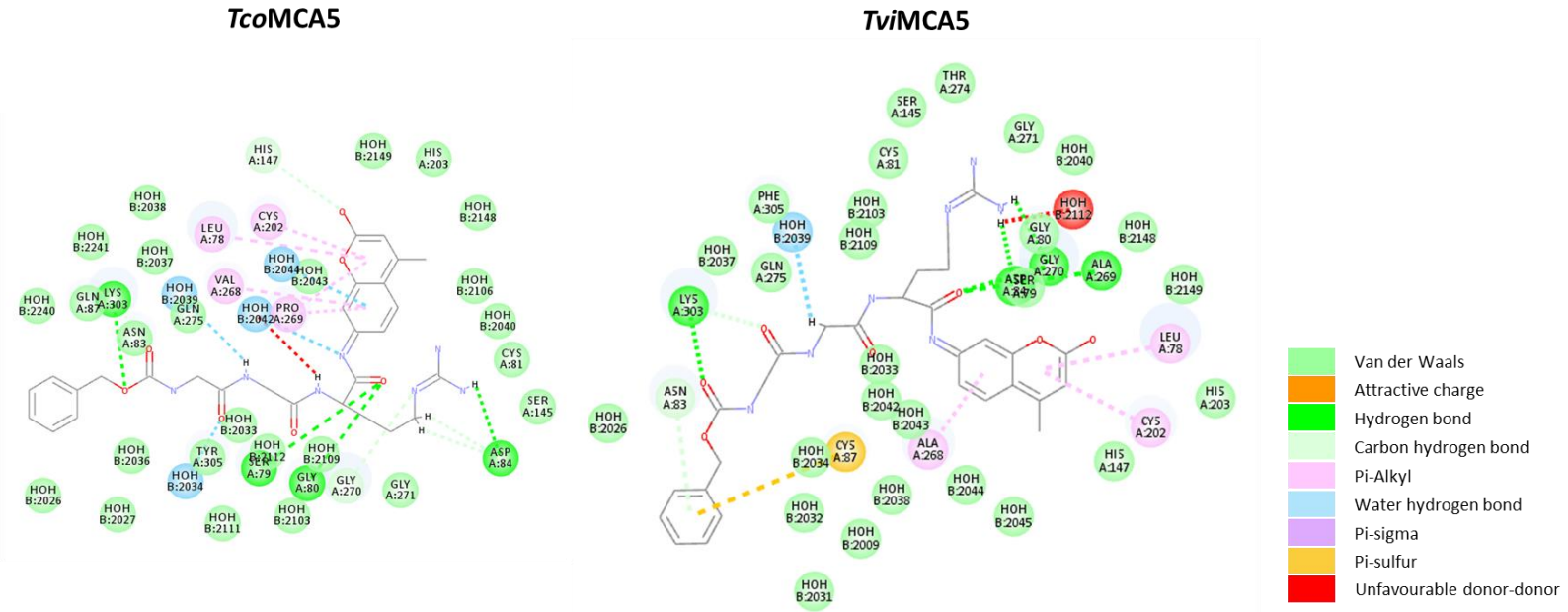


Appendix C5: The inhibitor, substrate and Berg ligand which had the highest binding affinity at the active site for *Tbb*MCA2 4AF8, prepared with 'all water'. The 2D interactions of the commercial inhibitor, antipain, commercial substrate, Z-Gly-Aly-Arg-AMC and Berg ligand, 777 docked in the active site of *Tbb*MCA2 4AF8.

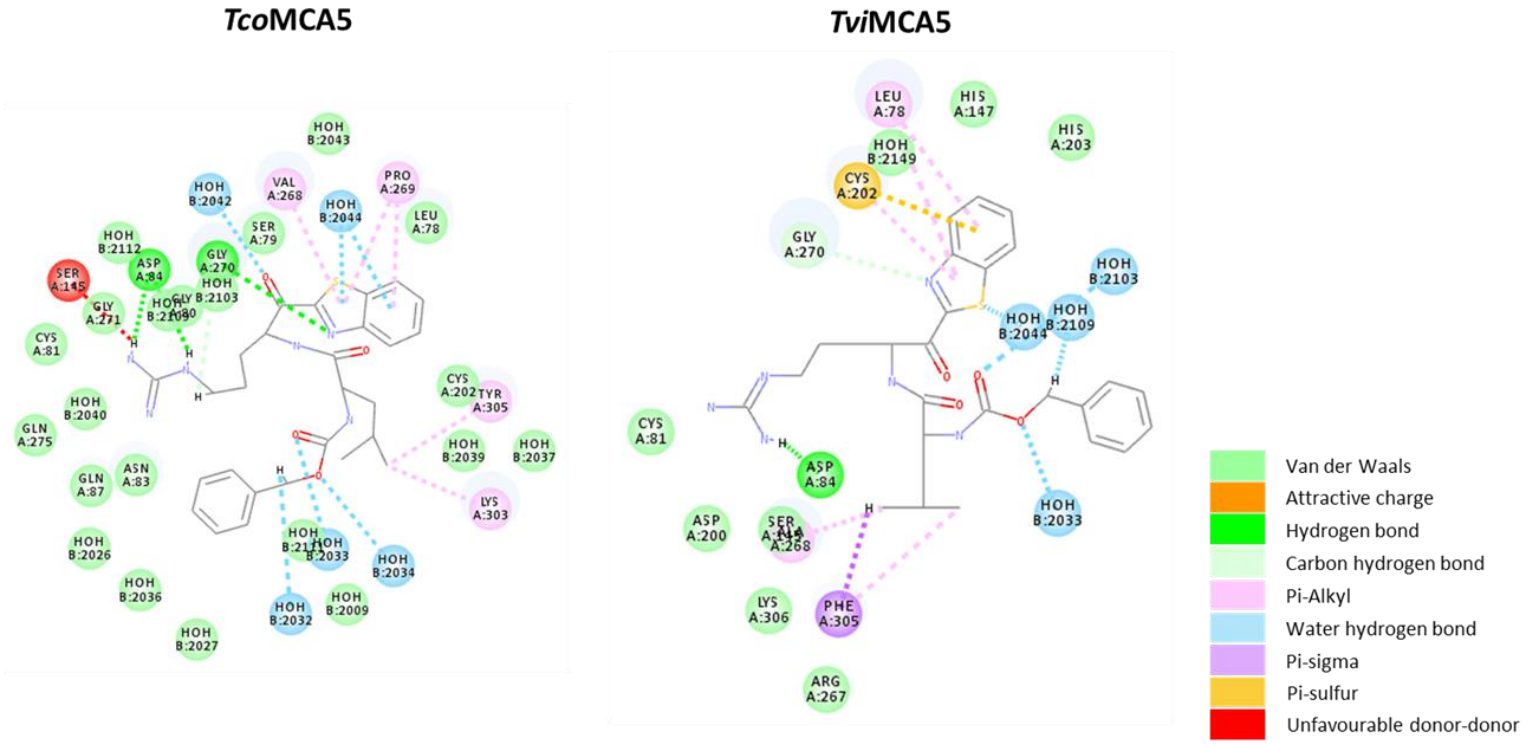
2D interactions of ligands docked in the active site of *Tco*MCA5 and *Tvi*MCA5



Appendix C6.1: The commercial inhibitor which had the highest binding affinity at the active site for *Tco*MCA5 and *Tvi*MCA5 prepared with 'all water'. The 2D interactions of the commercial inhibitor, antipain, docked in the active site of *Tco*MCA5 and *Tvi*MCA5.



Appendix C6.2: The commercial substrate which had the highest binding affinity at the active site for *Tco*MCA5 and *Tvi*MCA5 prepared with 'all water'. The 2D interactions of the commercial substrate Z-Gly-Gly-Arg-AMC, docked in the active site of *Tco*MCA5 and *Tvi*MCA5.

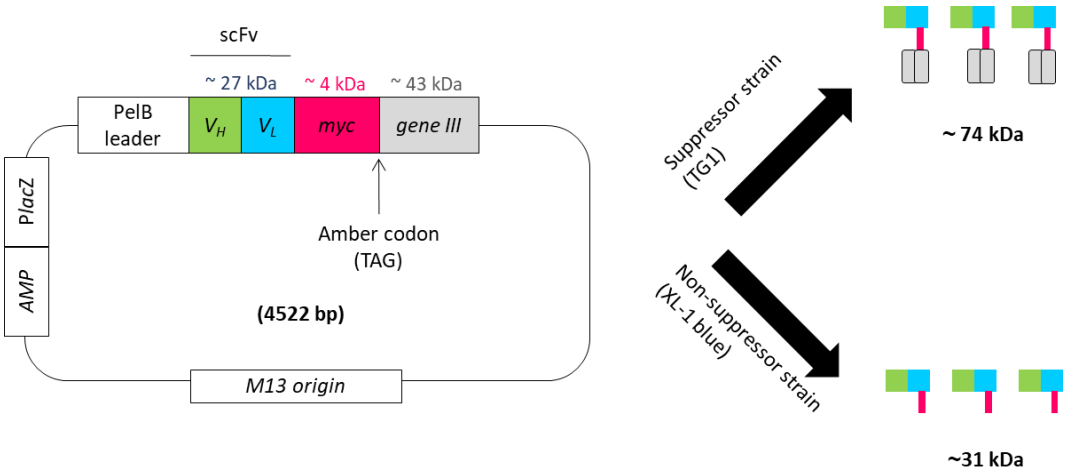


Appendix C6.3: The Berg ligand which had the highest binding affinity at the active site for *Tco*MCA5 and *Tvi*MCA5 prepared with 'all water'. The 2D interactions of the Berg ligand, 777, docked in the active site of *Tco*MCA5 and *Tvi*MCA5.

APPENDICES D

pHEN1 phagemid plasmid used for scFv production

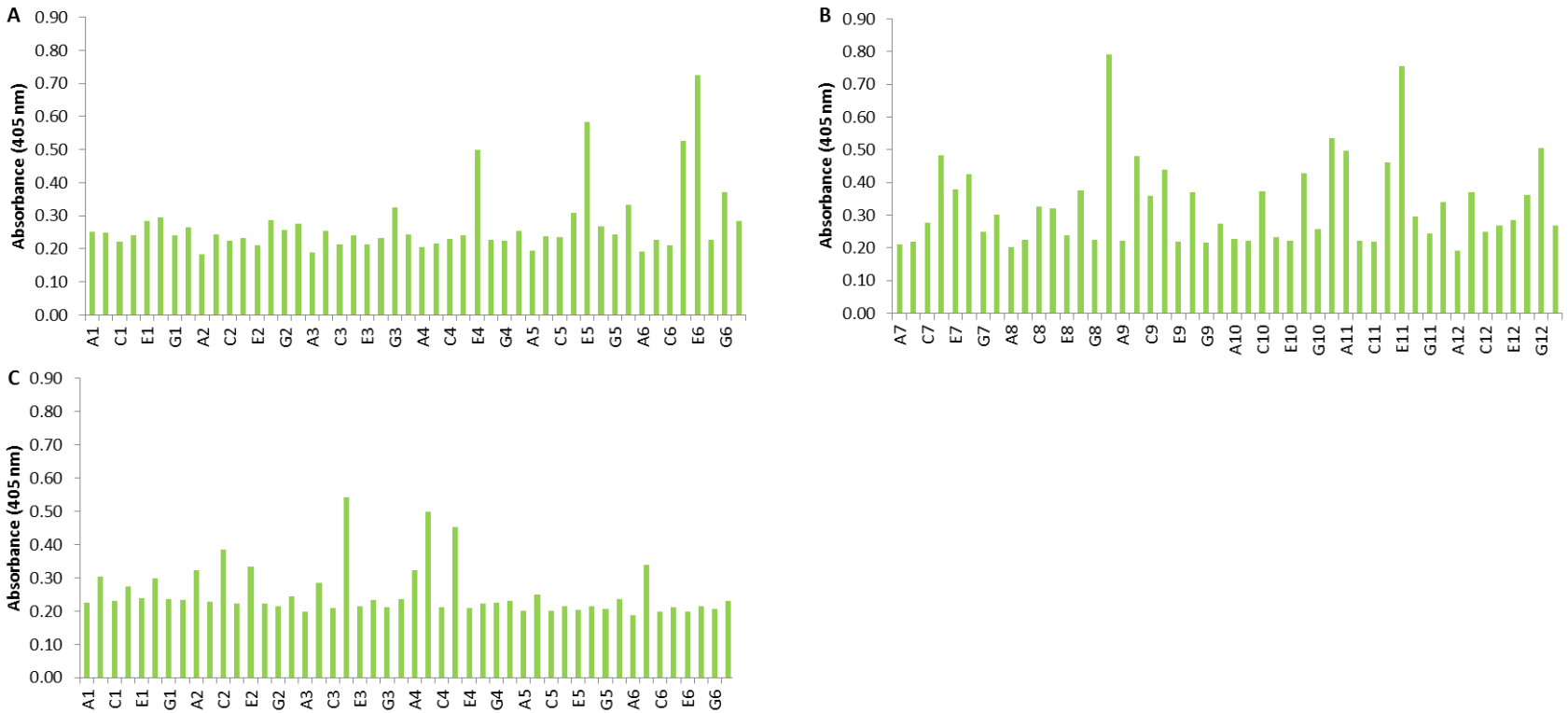
The map of the phagemid plasmid used in the *Nkuku*® phage display library is provided in Appendix D1, with the different resulting scFv formats and expected molecular weights when using suppressor and non-suppressor *E. coli* cells.



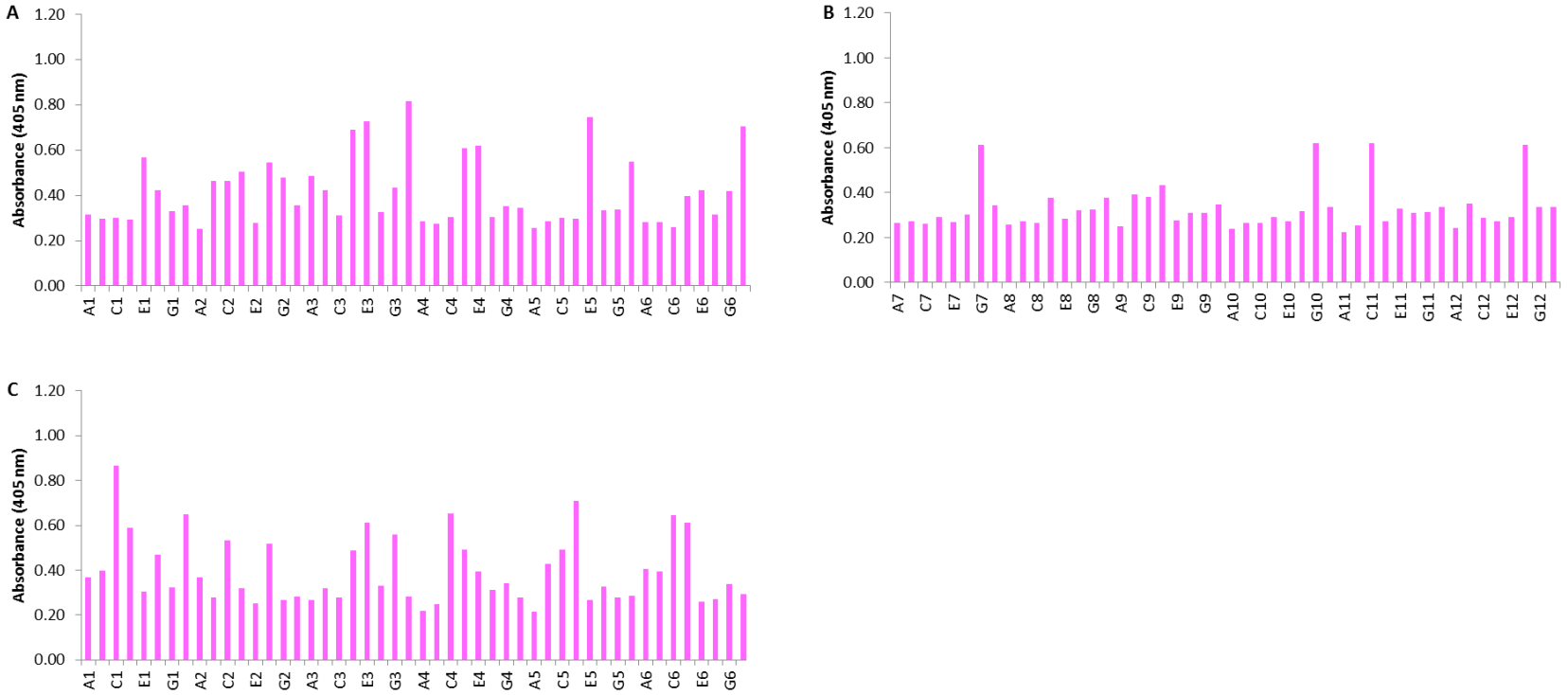
Appendix D1: Schematic diagram of the pHEN1 phagemid plasmid supplied in the *Nkuku*® phage display library. Adapted from Hoogenboom *et al*. (1991).

Phage binding and soluble scFv ELISAs

The phage binding ELISA format detects the pIII minor coat protein, which is fused to the myc-tagged scFv, when expressed in amber suppressor strains (Appendix D1). The soluble scFv ELISA format detected the myc tag using affinity purified chicken anti-myc IgY. This format is able to detect scFv from both amber suppressor and non-suppressor *E. coli* cells. The clones which had high absorbance values in both the phage binding and soluble scFv ELISA formats (Appendix D2 and D3) were used to define absorbance cut-off values, and noted in the decision matrix (Appendix D5).



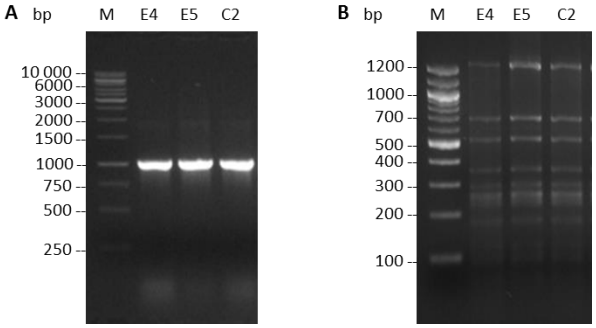
Appendix D2: Phage binding ELISA of *Tco*OPB and *Tvi*OPB panned phages. ELISA plates were coated with **(A and B)** *Tco*OPB and **(C)** *Tvi*OPB (1 µg/ml in PBS, pH 7.2), blocked with 3% BSA-PBS and incubated with equal volumes of expression supernatant of **(A)** *Tco*OPB pan 3, **(B)** *Tco*OPB pan 4 and **(C)** *Tvi*OPB pan 4 together with 6% (w/v) BSA-PBS. Thereafter, mouse anti-M13 (1:8000) was added. Goat anti-mouse IgG HRPO conjugate and ABTS·H_2O_2 were used as the detection system. Absorbance values are the result after 60 min of development.



Appendix D3: Soluble scFv ELISA of *Tco*OPB and *Tvi*OPB panned phages. ELISA plates were coated with **(A and B)** *Tco*OPB and **(C)** *Tvi*OPB (1 µg/ml in PBS, pH 7.2), blocked with 3% BSA-PBS and incubated with equal volumes of expression supernatant **(A)** *Tco*OPB pan 3, **(B)** *Tco*OPB pan 4 and **(C)** *Tvi*OPB pan 4 together with 6% (w/v) BSA-PBS. Thereafter, affinity purified chicken anti-myc IgY (0.5 µg/ml) was added. Rabbit anti-chicken IgY HRPO conjugate and ABTS·H_2O_2 were used as the detection system. Absorbance values are the result after 60 min of development.

Colony PCR and fingerprinting

In order to narrow down the selection of clones which would be sequenced, the plasmid DNA from the positive clones, identified from the two ELISA formats (Appendix D2 and D3), was isolated and subjected to colony PCR and DNA fingerprinting using AluI (Appendix D4). The clones which resulted in a 1 000 bp amplified product (panel A), were digested using AluI resulting in the same pattern of DNA fragments (panel B).



Appendix D4: Screening for full length scFv clones by PCR amplification and DNA fingerprinting. The isolated plasmid DNA from phage binding and soluble scFv ELISA positive clones were subjected to **(A)** PCR amplification of the scFv genes and **(B)** digestion with AluI. Samples were electrophoresed on 1.5% (w/v) and 3% (w/v) agarose gels respectively, both containing 0.5 µg/ml ethidium bromide. **(A)** M= O'GeneRuler™, **(B)** M= 100 bp ladder (NEB).

Decision matrix for the selection of scFv clones for expression

The results from the phage binding ELISA, soluble scFv ELISA and colony PCR, were used to generate a decision matrix to facilitate the selection of clones to be used for the expression of scFv (Appendix D5).

Appendix D5: Positive clones resulting from the various techniques to determine the best clones for scFv expression.

		cutoff	% positive[$]
*Tco*OPB pan3	Phage binding	0.4*	8
	Soluble scFv	0.5*	23
	colony PCR	1 000 bp[#]	100
*Tco*OPB pan 4	Phage binding	0.4*	19
	Soluble scFv	0.5*	8
	colony PCR	1 000 bp[#]	33
*Tvi*OPB pan 4	Phage binding	0.3*	8
	Soluble scFv	0.4*	12
	colony PCR	1 000 bp[#]	93

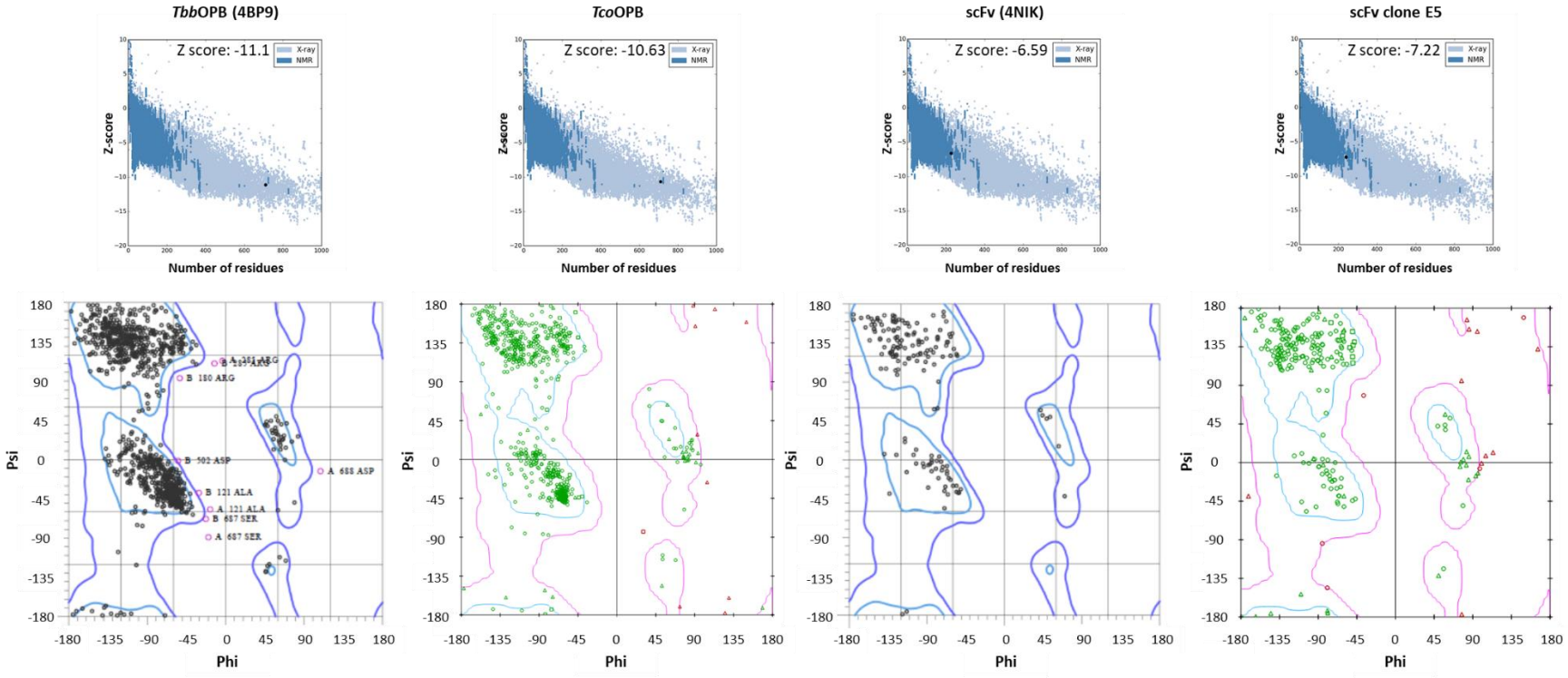
* The absorbance value where the cut off was made for the selection of positive clones.
# At 1 000 bp, a full length scFv was present. Anything below, would indicate a partial gene.
$ percentage positive was calculated by the division of the number of positive clones by the total number of clones tested and multiplied by 100.

Homology modelling and structure validation of modelled *Tco*OPB and clone E5 from *Tco*OPB pan 3

Enlarged sequence alignment of the *Tbb*OPB template and *Tco*OPB query.

Appendix D6: Homology modelling of *Tco*OPB. The alignment of the *Tco*OPB sequence to the *Tbb*OPB 4BP9 template is shown with the Qmean scores indicated in blue (high quality prediction) and red/orange (low quality prediction).

The validation of the generated structures, using the ProSA software (Wiederstein and Sippl, 2007) and the Ramachandran plots obtained from the Protein Data Bank (Berman *et al.*, 2000) and the Discovery studio software (Dassault Systèmes BIOVIA), of *Tco*OPB (Fig. 5.6) and scFv clone E5 modelled (Fig. 5.7) is reported in Appendix D7.



Appendix D7: Z-score and Ramachandran plots as an indication of model quality. The resulting Z-score plots of the individually uploaded PDB files, *Tbb*OPB (4BP9) and scFv (4NIK) along with the modelled *Tco*OPB and scFv clone E5 into the ProSA software (Wiederstein and Sippl, 2007). The Ramachandran plots for *Tbb*OPB 4BP9 (Canning *et al.*, 2013) and scFv 4NIK (Robin *et al.*, 2014) was obtained from the PDB (Berman *et al.*, 2000). Those for the *Tco*OPB and scFv clone E5 were generated using the Discovery Studio software (Dassault Systèmes BIOVIA).