IDENTIFICATION AND CHARACTERISATION OF NOVEL PATHOGENIC FACTORS OF *Trypanosoma congolense*

by

DAVITA PILLAY

B.Sc. (Hons) Biochemistry (UKZN)

Submitted in fulfillment of the academic requirements for the degree of Doctor of Philosophy in the Discipline of Biochemistry, School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal

Pietermaritzburg

December 2010

DECLARATION - PLAGIARISM

I, Davita Pillay, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis 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 thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed……

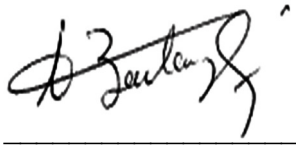
PREFACE

The experimental work described in this dissertation was carried out at the School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal, Pietermaritzburg, South Africa and the University of Bordeaux 2, France from August 2007 to December 2010, under the supervision of Prof. Theresa H.T. Coetzer and co-supervision of Dr Alain Boulangé.

These studies represent original work by the author and have not otherwise been submitted in any form for any degree or diploma to any University. Where use has been made of the work of others it is duly acknowledged in the text.

_____

Candidate: Davita Pillay



_____

Co-supervisor: Dr. Alain Boulangé

_____

Supervisor: Prof. Theresa H.T. Coetzer

ABSTRACT

Trypanosoma congolense is a major causative agent of the bovine disease trypanosomosis which has a considerable economic impact on sub-Saharan Africa. Current control methods for trypanosomosis are unsatisfactory and vaccine development has been hampered by antigenic variation.

An anti-disease vaccine is based on the idea that disease is caused by the pathogenic factors released by the parasite, rather than by the parasite itself. Therefore, if these pathogenic factors could be neutralised by antibodies produced by vaccination, the disease could be circumvented. The method used here for identification of novel pathogenic factors is based on the concept that trypanotolerant cattle are able to mitigate the disease by generating a specific immune response against a few key antigens (pathogenic factors).

Two immuno-affinity columns were therefore prepared: one containing IgG from non-infected sera and a second column containing IgG from trypanotolerant N'Dama cattle serially infected with *T. congolense.* The differential binding of antigens to the two columns allowed identification of antigens specifically recognised by the immune system of a trypanotolerant animal, i.e. potential pathogenic factors. The most promising antigens identified included several variant cathepsin L-like cysteine peptidases (CPs) and the Family M1 Clan MA aminopeptidases (APs). For the CPs, a study of the genetic organisation was conducted in order to further understand the variability present in this gene family.

To this end, two different mini-libraries of cathepsin L-like genes were prepared: one in which genes as different as possible from congopain (the major CP of *T. congolense*) were selected, and a second which contained all possible genes present in the congopain array. Analysis of the sequences obtained in these two mini-libraries showed that there was significant variability of the genes within the congopain array. Two variants of CPs, chosen for differences in their catalytic triads, were cloned for expression. The recombinantly expressed CP variants differed in substrate preferences from one another and from C2 (the recombinant truncated form of congopain), and surprisingly, all enzymes were active at physiological pH.

The two APs were cloned and expressed as insoluble inclusion bodies in an *E. coli* system, and subsequently refolded. The refolded APs showed a substrate preference for H-Ala-AMC, an optimum pH of 8.0, localisation to the cytoplasm and inhibition by puromycin. The two APs were not developmentally regulated and present in procyclic, metacyclic and bloodstream form parasites. Down-regulation of both APs by RNAi resulted in a slightly reduced growth rate in procyclic parasites *in vitro*. Immunisation of BALB/c mice with the APs did not provide protection when challenged with *T. congolense*. For an anti-disease vaccine to be protective, it would possibly have to include all pathogenic factors, including the two APs and at least one CP described in the present study.

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to the following people and institutions:

My supervisor and co-supervisor, Prof. Theresa Coetzer and Dr Alain Boulangé, for their guidance, patience and assistance throughout this study and for their critical reading of this thesis.

Prof. Théo Baltz, Virginie Coustou, Nicolas Biteau and Magali Thonnus for all their assistance, expertise and friendship, both in and out of the lab in Bordeaux, France.

Prof. Dean Goldring and Dr Edith Elliot for their innovative ideas and thought-provoking questions. Charmaine Ahrens and Robyn Hillebrand for their tireless help with administration over the years.

Liquid chromatography coupled to tandem mass spectrometry was conducted by Stéphane Claverol at the Plateforme Génomique and Fonctionnelle in Bordeaux, France. Two dimensional electrophoresis and blotting was carried out at the Institute de recherche pour le développement (IRD) and the Centre de cooperation internationale en recherché agronomique pour le développement (CIRAD, Montpellier, France) with the kind assistance of Jean-Paul Brizard and Phillipe Holzmuller.

This work was supported by the Commission of the European Communities' Sixth Framework Programme, priority INCO-DEV, project Trypadvac2, contract PL003716, the PROTEA/PICS project funded by the NRF (South Africa) and the CNRS (France) and facilitated by M. Vincent Baron at the French Embassy in South Africa.

My colleagues Phillia Vukea, Richard Kangethe, Hlumani Ndlovu, Lorelle Bizaaré, Cara Bartlett, Perina Vather, Sabelo Hadebe, Ike Achilonu, Jackie Viljoen, Bridgette Cumming, Dave Choveaux and Laura Huson for their company and advice.

My parents, Sanjay and Manu Pillay, for their support. My brother, Adushan because I have to.

Lastly, all my cats, past and present, for keeping me sane.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2x YT	yeast tryptone media
AAT	African animal trypanosomosis
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride
AMC	aminomethyl coumarin
AMT	acetate-MES-tris buffer
AP(s)	aminopeptidase(s)
BBB	blood brain barrier
BCA	bicinchoninic acid
BLASTN	basic local alignment search tool - nucleotide
BMGY	buffered media glycerol yeast extract
BMM	buffered minimal media
BSA	bovine serum albumin
CAF-T	complement activating factor
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CP(s)	cysteine peptidase(s)
col. vols	column volumes
DEAE	diethyl-aminoethyl
DMF	dimethyl formamide
DMK	dimethyl ketone
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
E-64	*trans*-epoxysuccinyl-L-leucyl-amido(4-guanidino)butane
ECL	enhanced chemoluminescence
EDTA	ethylene diamine tetra-acetic acid
ELISA	enzyme-linked immunosorbent assay
GSH	reduced glutathione
GSSG	oxidised glutathione
HRPO	horse radish peroxidase
i.p.	intra-peritoneal
IEC	ion exchange chromatography
IEF	iso-electric focusing
IFA	immunofluorescence assay
IFNγ	interferon gamma
Ig	immunoglobulin

ILRI	International Livestock Research Institute
IPG	immobilised pH gradient
LAP(s)	leucyl aminopeptidase(s)
LC-MS/MS	liquid chromatography coupled to tandem mass spectrometry
mAb	monoclonal antibody
MALDI-ToF	matrix-assisted laser desorption/ionisation – time of flight
MAP p15	microtubule associated protein 15
MBS	*m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester
MEC	molecular exclusion chromatography
MS	mass spectrometry
OD_{600}	optical density at 600 nm
OpdB	oligopeptididase B
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PCV	packed cell volume
PEG	polyethylene glycol
PSG	phosphate saline glucose
RBC(s)	red blood cell(s)
RT	room temperature (approximately 25°C)
SAP	shrimp alkaline phosphatase
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SERA	serine repeat antigens
sVSG	soluble variant surface glycoprotein
TBS	tris buffered saline
TCA	trichloroacetate
TNFα	tumour necrosis factor alpha
TPP	three phase partitioning
VAT	variant antigen type
VSG	variant surface glycoprotein
WT	wild type
YPD	yeast peptone dextrose media
Z	benzyloxycarbonyl

CHAPTER 1: LITERATURE REVIEW

1.1 TRYPANOSOMES

In 1894, David Bruce first linked trypanosomosis to the presence of protozoan parasites in the bloodstream while investigating the death of cattle in Zululand, South Africa (Duggan, 1977; Brown, 2008). Two African trypanosome sub-species are pathogenic to humans, namely, *Trypanosoma brucei rhodesiense* and *T. b. gambiense,* which cause the disease sleeping sickness or human African trypanosomosis. *T. congolense*, *T. vivax* and *T. b. brucei* are all transmitted by the tsetse fly and are the causative agents of African animal trypanosomosis (AAT) in cattle in sub-Saharan Africa. These species also infect other domestic animals, including camels, horses, dogs, sheep, goats and pigs (Steverding, 2008). *T. vivax* also infects ungulates in South America, where it is transmitted non-cyclically by hematophagous flies rather than by tsetse flies (Osório *et al.*, 2008). *T. evansi*, transmitted by tabanids, is the causative agent of the disease *surra,* prevalent in Africa, South America and Asia (Brun *et al.*, 1998). *T. equiperdum* has evolved a sexual mode of transmission and is transmitted during coitus in horses and donkeys causing the disease dourine (Claes *et al.*, 2005). *T. cruzi*, the causative agent of Chagas disease in South America, is an intracellular parasite spread by biting flies (Tanowitz et al., 1992; Rassi et al., 2010).

1.1.1 Classification

African trypanosomes are flagellated haemoprotozoan parasites belonging to the order Kinetoplastida, which are characterised by the presence of a rod-shaped DNA-containing kinetoplast; by a single flagellum attached to the body as an undulating membrane; and by a single, highly branched mitochondrion (Uilenberg, 1998). David Bruce first classified pathogenic African trypanosomes into three groups viz. the *T. brucei* group, the *T. vivax* group and the pig infective trypanosome, *T. simiae* (Bruce, 1914). Bruce entirely omitted *T. congolense* from his classification, a species which is now considered to be one of the most significant causes of AAT. Molecular methods of characterisation have now superseded the traditional (and fallible) morphological analysis as means of discriminating between species. As a result, at least four new sub-types and one new species (*T. godfreyi*) have been described in the last thirty years (Adams *et al.*, 2010). Figure 1.1 shows a recent classification of trypanosomes of medical and veterinary importance.

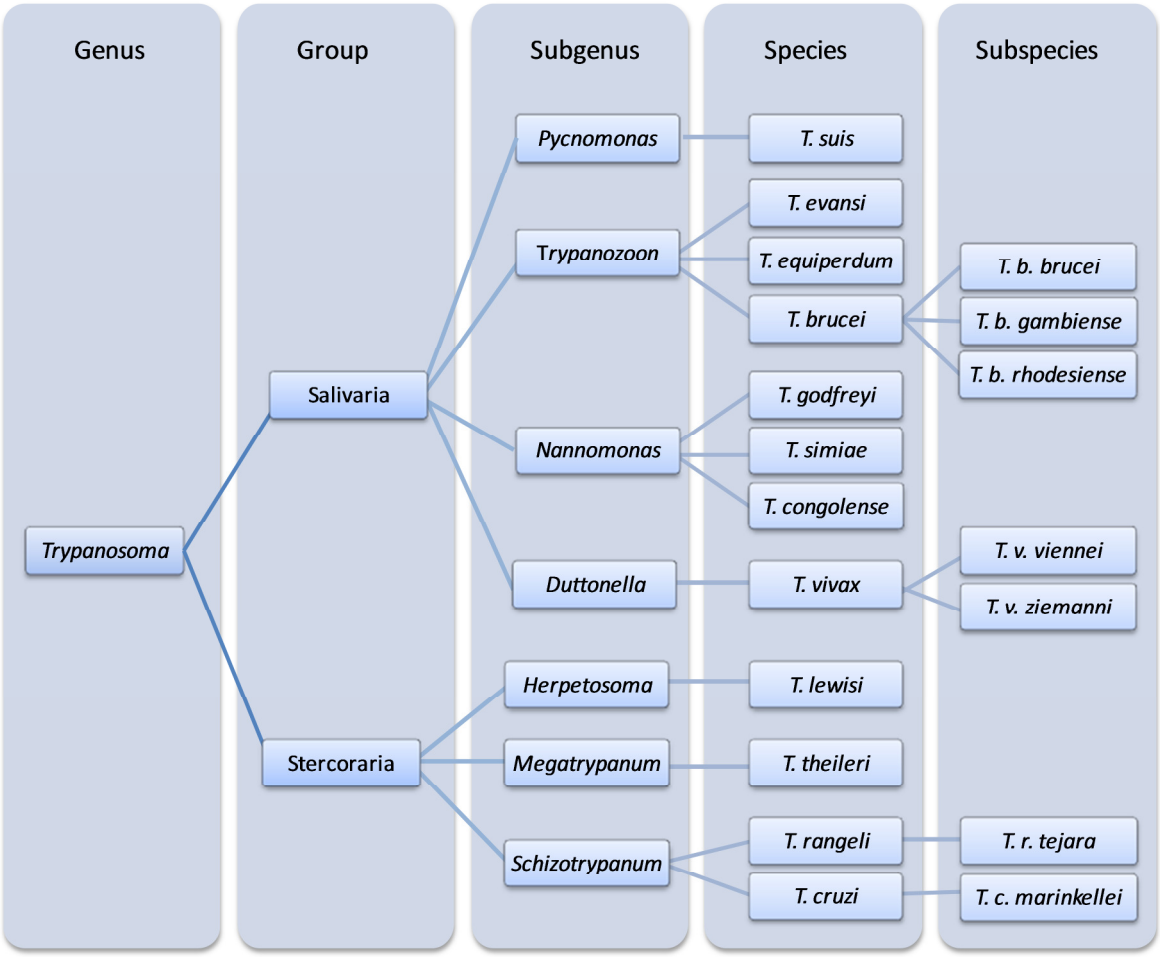


Figure 1.1 Classification of trypanosomes of medical and veterinary significance (adapted from Stevens and Brisse, 2004)

In 1924, researchers Lloyd and Johnson, divided mammalian infective trypanosomes into two groups according to which part of the vector the parasite develops into the infective metacyclic form (reviewed in Enyaru *et al.*, 2010). In the group *Stercoraria*, transmission occurs by contamination of broken skin by the insect faeces; while in *Salivaria*, trypanosomes are injected into the mammalian host from the buccal part of the vector during the bloodmeal. Interestingly, antigenic variation, the immune evasion mechanism (see Section 1.1.2.3) evolved independently in Salivarian trypanosomes after the evolutionary split between the two groups (Haag *et al.*, 1998). There are, at present, four subgenera of Salivarian trypanosomes, viz. *Duttonella, Nannomonas, Trypanozoon* and *Pycnomonas* (Figure 1.1). The three tsetse fly-transmitted trypanosome species responsible for AAT in cattle: *T. vivax*, *T. congolense* and *T. brucei*, belong to the subgenera *Duttonella*, *Nannomonas* and *Trypanozoon* respectively (Stevens and Brisse, 2004).

Three distinct species occur within the *Nannomonas* subgenus, viz. *T. congolense,* which is the most common agent of AAT; *T. godfreyi* and *T. simiae*, which are both found in wild and domestic animals in sub-Saharan Africa (Majiwa and Webster, 1987; Majiwa et al., 1993; McNamara et al., 1994). These three species are difficult to distinguish from each other due to morphological similarity. There is, however, sufficient genetic diversity between them at molecular level, specifically the 18S RNA sequences, to differentiate between these three *Nannomonas* species (Gibson, 2007). Furthermore, in the last 30 years it was found that there are three recognised types of *T. congolense* which appear morphologically identical, but are distinguished based on the habitat of the vector; viz. savannah, riverine/forest, and Kenyan coast (Kilifi) types. These three types of *T. congolense* can also be differentiated by PCR based on the variations in the major satellite DNA repeat (Adams *et al*., 2010).

1.1.2 Biology of African trypanosomes

1.1.2.1 Morphology

African trypanosomes are unicellular organisms and may vary in size from 9 to over 25 µm. The parasites are streamlined in shape and tapered at both ends to assist with movement in the bloodstream (Figure 1.2). The pellicle, a thin cell layer supporting the cell membrane, allows enough flexibility for movement while being rigid enough to retain the tapered shape (Vickerman, 1969b). The surface of Salivarian trypanosomes (with the exception of the flagellar pocket) is covered by a glycoprotein coat, called the variant surface glycoprotein (VSG) which is essential for immune evasion by the mechanism of antigenic variation (see Section 1.1.2.3) (Vickerman, 1969b).

A single flagellum, produced by a parabasal body, and an undulating membrane (usually inconspicuous in *T. congolense*), produced by the pellicle and the cytoplasm, are both essential for movement of the parasite through the bloodstream (Bastin *et al*., 2000b; Broadhead *et al*., 2006). In addition, the flagellum of *T. brucei* has been shown to play a significant role in morphogenesis, cell division and immune evasion, since the flagellum is involved in attachment to host surfaces (Ralston and Hill, 2008). Interestingly, African trypanosomes have been shown to exhibit a phenomenon known as “social motility” whereby groups of trypanosomes migrate in specific patterns which are distinctly different from the movement of single parasites. This behaviour is thought to be coordinated by similar signalling factors which influence host-parasite interactions, as well as being mediated by the flagellum (Oberholzer *et al*., 2010).

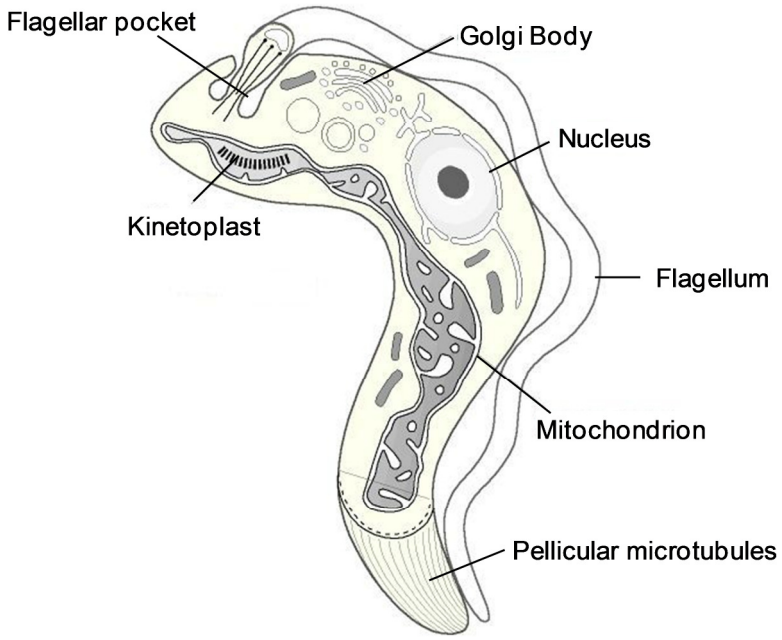


Figure 1.2 Diagrammatic representation of *Trypanosoma congolense* indicating the most significant morphological features (adapted from Vickerman, 1969a).

T. congolense lacks a free flagellum (Figure 1.2) and this is used in differential diagnosis since both *T. brucei* and *T.* vivax do possess a free flagellum (Osório *et al.*, 2008). The point of attachment of the flagellum to the body of the trypanosome is via an invagination of the plasma membrane known as the flagellar pocket (Figure 1.2). The flagellar pocket region is the sole site of endocytosis of the trypanosome, as well as being involved in protein trafficking, i.e. the post-translational modifications and recycling of cellular proteins. The trypanosome's ability to evade the immune system is also dependent on the protein trafficking that occurs at the flagellar pocket, since recycling of the VSG coat and degradation of immune complexes bound to these VSG, occurs only at the flagellar pocket (Field and Carrington, 2009).

The position of the nucleus differs between species, being centrally situated in *T. congolense* and located towards the back in *T. brucei*. The kinetoplast, of medium size in *T. congolense*, but larger in *T. vivax*, was shown to be necessary for division, metabolism and cyclical transmission by tsetse flies (Lun and Desser, 1995). The large kinetoplast of *T. vivax* (approximately 1.1 µm) is useful for differential diagnosis (Osório *et al.*, 2008). The kinetoplast is usually found adjacent to the parabasal body from which the flagellum arises (Vickerman *et al.*, 1988).

1.1.2.2 Genomic organisation

Two genomes are present in African trypanosomes, one within the nucleus and one within the kinetoplast. Kinetoplast DNA is condensed into mini-circles (~1 kb) and maxi-circles (~22 kb). Maxi-circles are typically found in low copy number (~12/cell) and have conserved sequences encoding ribosomal RNA and some mitochondrial proteins. Mini-circles occur in high copy number (5000-10 000/cell) and are composed of highly variable sequences (Liu *et al.*, 2005).

The nuclear genome is approximately 3.5 x 10^7 bp in length, roughly the same as the genomes of parasites such as *Plasmodium falciparum, Leishmania major* and *T. cruzi*. The nuclear genome is of low complexity since 68% consists of single copy sequences. Most of the nuclear genome is diploid, with some exceptions including the VSG genes (see Section 1.1.2.3), VSG expression sites and mini-chromosomes (Kooy et al., 1989; El-Sayed et al., 2000). It is important to note that the sequence information regarding the *T. congolense* genome is based upon the *T. brucei* genome sequence (Sanger and Coulson, 1975; Melville et al., 2004).

1.1.2.3 Antigenic variation

Trypanosomes have evolved the process of antigenic variation to evade the host immune system. With the exception of the flagellar pocket, the outer surface of the bloodstream forms of trypanosomes is entirely coated with VSG (Vickerman, 1969b). Each parasite in a population has the genetic potential to produce an antigenically distinct VSG (Taylor and Rudenko, 2006). The different VSG coats are termed variant antigen types (VATs). Antigenic variation is the term used to describe the varying expression of these VSG coats, resulting in antigenically distinct populations of trypanosomes of each VAT present in the blood over time. As a specific immune response develops against one VAT, the number of trypanosomes with homologous VAT declines; those trypanosomes with a heterologous VAT will survive, divide and initiate the next wave of parasitaemia (Barry and Carrington, 2004).

The expression of different VSG coats is thought to occur when a VSG gene, generally located in a non-transcribed region (usually sub-telomeric parts of the mini-chromosomes), is duplicated into a transcriptionally active region (Barry *et al.*, 2005). Therefore, at a molecular level, this process is based on the switching of expression of various VSG gene segments in a similar manner to which immunoglobulin populations vary in higher eukaryotes (Marcello and Barry, 2007).

The VSG proteins are a major immune target and anti-VSG antibodies have been shown to limit parasitaemia (Guirnalda *et al.*, 2007). However, cattle immunised with purified VSGs are protected against homologous, but not heterologous, infections (Taylor, 1998). This phenomenon of antigenic variation has greatly hindered the development of a conventional vaccine against trypanosomosis (Zambrano-Villa *et al.*, 2002).

1.1.2.4 Life cycle of African trypanosomes

David Livingstone first proposed the link between tsetse flies and AAT when he witnessed cattle dying after being bitten by tsetse in the Lake Nyasa (now Malawi) area (reviewed in Steverding, 2008). At present, there are 23 different species and eight recognised subspecies of the *Glossina* genus which is divided into three groups: *fusca* (forest), *palpalis* (riverine), and *morsitans* (savannah) (Krafsur, 2009). Both male and female flies are able to carry trypanosomes and infect the mammalian host during feeding.

T. congolense, T. vivax and *T. brucei* all spend part of their life cycle in the tsetse fly vector and part in a mammalian host (Figure 1.3). The infective non-dividing metacyclic forms of the parasite are found in the mouth parts of the tsetse fly and are transmitted to the mammalian host when the fly bites an animal. During the blood meal, the fly injects saliva into the bite to prevent coagulation of the blood. The infective trypanosomes initially divide at the site of infection, often forming a chancre, and are released into the bloodstream via lymph vessels and lymph nodes (Vickerman et al., 1988; Uilenberg, 1998).

Once in the bloodstream, *T. brucei* differentiates into rapidly dividing, long, slender, bloodstream forms (trypomastigotes) which spread to different organs of the host via the bloodstream (Vickerman *et al.*, 1988). Trypomastigotes divide by binary fission and subsequently differentiate into short, stumpy forms of the parasite which are ingested by a tsetse fly feeding on the infected animal (Matthews *et al.*, 2004).

After being ingested by the fly, the parasite differentiates into a rapidly dividing procyclic form in the midgut of the fly. As the parasite changes into the procyclic forms, it alters its enzyme systems to utilise different nutrients at different temperatures. One example of this is that *T. brucei* develops a functional mitochondrion to provide energy via oxidative phosphorylation in addition to the usual glycolysis of the bloodstream form (Sharma *et al.*, 2009).

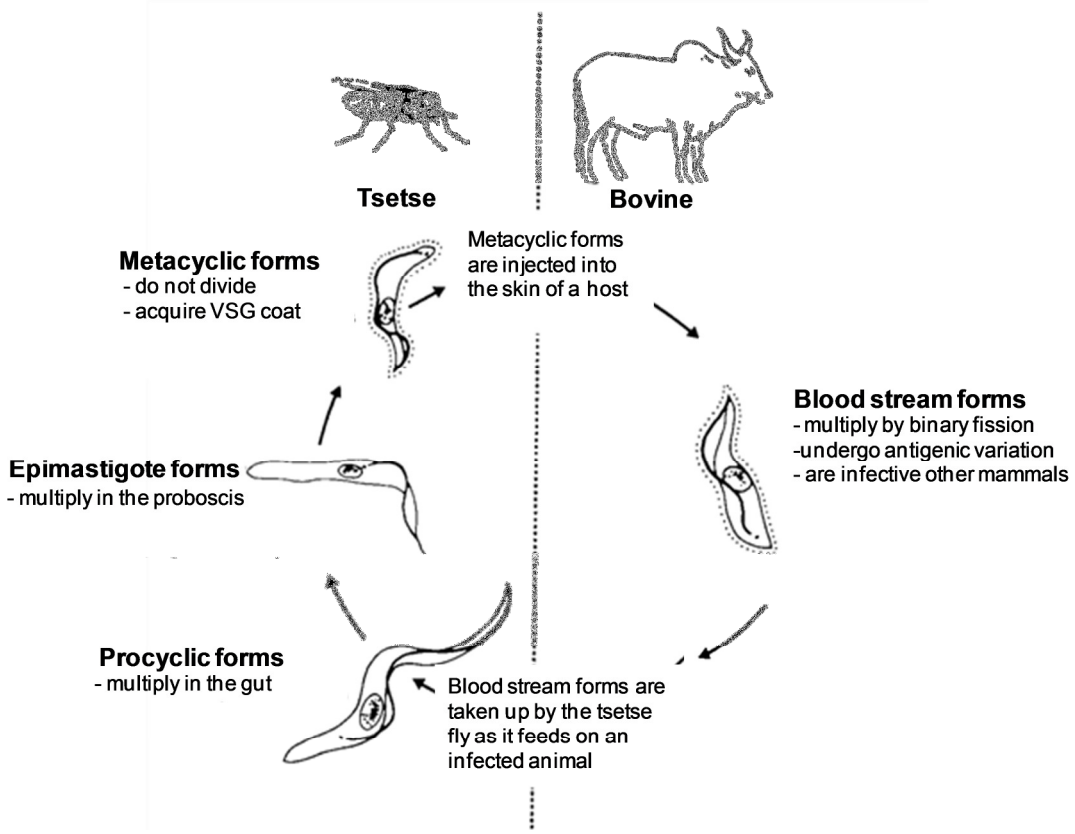


Figure 1.3 Life cycle of *Trypanosoma congolense* (adapted from Authié, 1994).

Approximately three days after ingestion, parasites begin to die, leading to the complete elimination of infection in some flies. In flies where the parasite does manage to establish an infection, it invades the ectoperitrophic space, and eventually moves into the proventriculus region (Aksoy, 2003). The procyclic forms of *T. brucei* subsequently differentiate into the epimastigote form in the salivary glands, by the shedding of their VSG coat. *T. congolense* develops into procyclic forms in the proboscis of the fly. *T. vivax,* however, develops entirely in the proboscis of the vector with no developmental stages occurring in the midgut or salivary glands of the fly (Matthews *et al.*, 2004).

The cycle repeats as the epimastigote forms change, once again, into the infective short metacyclic form (Englund *et al.*, 1982), hence the term "cyclical transmission". The period between ingestion of the trypomastigotes from the mammalian host and the change into the infective metacyclic form varies from one to three weeks. It has long been thought that trypanosomes exchange genetic material by meiosis and mating, although the exact point in the life cycle at which this occurs was never clear (Vickerman, 1986). However, it has recently been definitively shown that genetic exchange does occur in *T. brucei* in the salivary glands of the insect vector (Gibson *et al.*, 2008; Hide, 2008) and in *T. congolense* (Holzmuller *et al.*, 2010).

1.2 AFRICAN ANIMAL TRYPANOSOMOSIS

African animal trypanosomosis (AAT) is of considerable economic importance, causing huge losses as a result of cattle mortality and morbidity, diagnosis and treatment costs and decreases in meat and milk production due to unhealthy animals. Few large scale studies have been done to evaluate the actual cost of AAT. The most comprehensive of these, conducted by Kristjanson and colleagues (1999) estimates the losses due to AAT at USD $ 1 - 5 billion per year in Africa. A third of the area of the African continent (approximately 10 million km²) is infested by tsetse flies and 60 million cattle across 37 countries are at risk of trypanosomosis (Kristjanson *et al.*, 1999) (Figure 1.4).

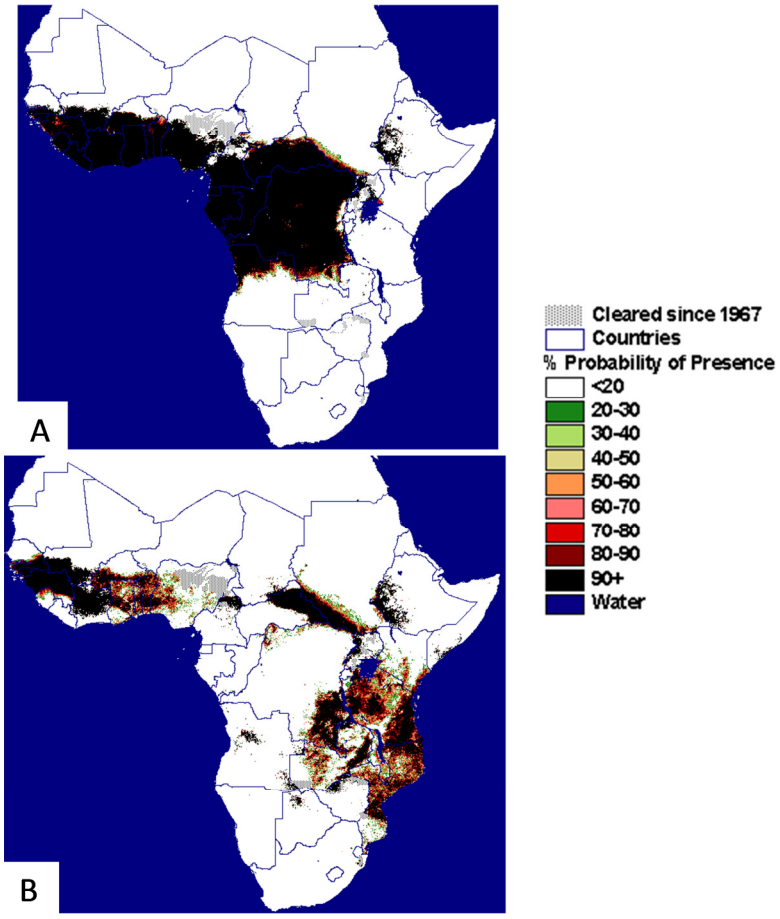


Figure 1.4 Distribution of (A) Palpalis and (B) Morsitans group tsetse flies predicted from satellite imagery (Wint and Rogers, 1999)

The distribution of trypanosomes depends on the geographic distribution of the tsetse fly vector (Van den Bossche, 2001). It must be noted that *T. vivax* can be transmitted by biting flies and is, therefore, present in South America although no tsetse flies are found there (Osório *et al.*, 2008).

Several approaches have been employed to control AAT, including the use of trypanocides, controlling the tsetse fly populations and the use of livestock which are innately resistant to the disease (trypanotolerant) (d'Ieteren et al., 1998; Anene et al., 2001; Gooding and Krafsur, 2005; d'Ieteren and Kimani, 2007). The trypanocidal drugs which have been available for the past fifty years are costly, and drug resistant forms of the parasites have been found in Africa (Anene *et al.*, 2001; Maser *et al.*, 2003). The control of tsetse fly populations by destruction of their habitat and the use of insecticides have been effective but are not ecologically sustainable (Torr *et al.*, 2005).

1.2.1 Pathogenesis

Trypanosomes are responsible for three known pathological states in African livestock, which have been given the common names *nagana* (bovine trypanosomosis), surra (caused by *T. evansi*), and dourine (caused by *T. equiperdum*) (Table 1.1). Bovine trypanosomosis is caused mainly by the tsetse transmitted parasites *T. congolense, T. vivax,* and *T. b. brucei* (Stevens and Brisse, 2004).

Table 1.1 Overview of the pathogenic salivarian trypanosomes (adapted from Gibson, 2007).

Subgenus	Species	Disease	Distribution	Potential mammalian hosts
Trypanozoon	*T. brucei brucei*	Nagana	Tropical Africa	Wild and domestic animals
	T. brucei rhodesiense	Rhodesian sleeping sickness	East Africa	Humans, wild and domestic animals
	T. brucei gambiense	Gambian sleeping sickness	West and Central Africa	
	T. evansi	Surra	North Africa, Middle East, Asia, South America	Wild and domestic animals
	T. equiperdum	Dourine		Equines
Nannomonas	*T. congolense*	Nagana	Tropical Africa	Wild and domestic animals
	T. simiae	Acute porcine trypanosomosis	Tropical Africa South America	
Duttonella	*T. vivax*	Nagana	Tropical Africa	Wild and domestic animals
Pycnomonas	*T. suis*	Porcine trypanosomosis	Tropical Africa	Suids

These pathogenic trypanosomes can be separated into two categories: "haematic" trypanosomes, including *T. congolense* which are limited to the blood vessels of the host animal and "humoral" trypanosomes including *T. b. brucei, T. b. rhodesiense* and *T. b. gambiense* which are present in plasma, intracellular tissues, body cavity fluids and are able to cross the blood brain barrier (Uilenberg, 1998; Lonsdale-Eccles and Grab, 2002). Although *T. b. brucei* is capable of tissue invasion, it is generally less pathogenic

than *T. congolense* and *T. vivax.* Anaemia is a symptom typical of infections of both groups, while infection by "humoral" trypanosomes may result in necrosis and inflammation (Losos and Ikede, 1970; Gibson, 2007).

The progress of bovine trypanosomosis is dependent on the animal host, and trypanosome species and strain. However, some clinical signs of the disease are common to most domestic animals, irrespective of the species of trypanosome involved. Detection of parasitaemia and onset of anaemia occur one to three weeks after infection depending on the virulence of the strain, infective dose and immune status of the host animal. The initial chancre is caused by the rapid multiplication of trypanosomes in the skin in the area around the tsetse fly bite, and is visible two to three weeks after infection. The parasites subsequently enter the lymph nodes surrounding the region of the bite, resulting in the enlargement of the lymph nodes (Mattioli *et al.*, 1999). Once in the lymphatic system, the parasites have access to the host bloodstream, and consequently, blood filtering organs such as the spleen and liver (Taylor and Authié, 2004).

Initial clinical signs of the acute phase of the disease include enlarged lymph nodes and spleen, weakness, lethargy, general loss of condition, abortion and reduced milk production. Infection may also result in damage to the bone marrow, heart, endocrine glands, and reproductive glands and, in the case of *T. brucei*, the central nervous system. Death may occur in the first few weeks, or the animal may stabilise and begin a slow recovery. Animals most likely progress to the chronic phase of the disease characterised by intermittent parasitaemia, stunting, wasting and infertility (Bengaly *et al.*, 2002). Chronic trypanosomosis usually results in death by congestive heart failure, which is a result of prolonged anaemia, myocardial damage and increased vascular permeability (Taylor and Authié, 2004).

1.2.2 Diagnosis

Diagnosis of AAT can be performed by either clinical diagnosis, parasitological methods, which involve visualising the parasite itself, or by indirect methods which detect either compounds of the parasite, or specific antibodies directed against parasite antigens.

Clinical diagnosis is usually used for AAT simply because under field conditions, physical examination of the animal is often the sole means of diagnosis available (Eisler *et al.*, 2004). However, none of the symptoms displayed by trypanosome infected animals (see Section 1.2.2) are uniquely characteristic, and, as a result, may be mistaken for one of many other tropical diseases.

Direct parasitological methods usually involve microscopy-based diagnosis by making a film of either fresh blood or lymph and identifying the parasite morphologically. A major disadvantage of this technique is the low degree of sensitivity, particularly during the chronic phase of the disease. The sensitivity may be increased by concentrating the trypanosomes by centrifugation, known as the buffy coat technique (Murray *et al.*, 1977; Paris *et al.*, 1982). A modification of this method includes the removal of blood cells from the preparation by use of anion exchange chromatography (Nantulya, 1990). Also, the high degree of morphological similarity between different species of trypanosomes often makes differential diagnosis difficult. These are, however, the most simple and inexpensive methods of diagnosing AAT (Uilenberg, 1998).

Indirect methods for diagnosis could be serological or molecular tests. Serological tests include the indirect immunofluorescent antibody (IFA) test, indirect and direct enzyme-linked immunosorbent assays (ELISAs) and the complement fixation test. In an IFA test, sera are allowed to interact with parasite antigens (usually whole lysate) fixed onto a microscope slide using glutaraldehyde or formalin (Wilson, 1969; Chandler and Watts, 1988). Positive reactions are detected by a suitable antibody conjugated to a fluorescent substrate, usually fluoresceine. The IFA test is not species-specific since cross-reactions occur between *T. congolense, T. brucei,* and *T. vivax.* Also, the high cost of a fluorescence microscope is a disadvantage (Eisler *et al.*, 2004).

The indirect and direct ELISAs used for diagnosis of AAT can be species-specific if specific antigens are used (Desquesnes and Dávila, 2002). An example of a trypanosome specific ELISA is the test based on the *T. congolense* mitochondrial heat shock protein 70 (Bannai *et al.*, 2003) and the related immunoglobulin binding protein (heat shock protein 70 or BiP) (Boulangé *et al.*, 2002). Both these proteins have been proposed as diagnostic antigens due to the specificity of the ELISAs as well as the relative ease of procuring sufficient antigen for the test using recombinant expression. A more recent improvement on this ELISA was the use of a monoclonal antibody against recombinant BiP in an inhibition ELISA (Bossard *et al.*, 2010). This inhibition ELISA was shown to specifically identify *T. congolense* and *T. vivax* infections (*T. brucei* antibodies were not well recognised by the ELISA). Significantly, this inhibition ELISA was able to differentiate between currently sick animals and those which had been cured over three months prior to the test (Bossard *et al.*, 2010). Another advantage of ELISA assays is

that automation is possible, enabling the processing of large numbers of samples (Cabrera *et al.*, 2009).

The complement fixation test for animal trypanosomosis was developed by Edward Watson, who was eventually credited with the elimination of dourine in horses from Canada in 1920 due to the early detection possible with this test (reviewed in Derbyshire and Nielsen, 1997). Due to cross-reactivity and problems with preparation of pure antigen, the complement fixation test is still only useful for the diagnosis of dourine in horses and not other animal infective trypanosomes (Luckins, 1992).

The two main types of molecular tests are labelled DNA probes and polymerase chain reaction (PCR). DNA probes may be radioactively labelled, which is not suitable for use in field laboratories, and enzyme-labelled probes tend to have a much reduced sensitivity. PCR allows species-specific detection of parasites and is more sensitive than radio-labelled probes (Desquesnes and Dávila, 2002). More recently, a PCR-ELISA has been developed for the detection of *T. congolense* (Sow *et al.*, 2006). The combination of these two techniques minimises the cross-reactions inherent in ELISAs as well as eliminating the problem of differentiating between sick and cured animals which all serological tests are prone to. Furthermore, the PCR-ELISA is a much more powerful diagnostic technique that the buffy coat technique developed by Murray *et al.* (1977) which is currently the most widely used technique for diagnosis of AAT. This PCR-ELISA has recently been expanded to differentiate between *T. congolense* (including the riverine/forest, Kilifi and savannah subgroups), *T. vivax* and the *Trypanozoon* subspecies (including *T. brucei spp.*, *T. evansi* and *T. equiperdum*) by PCR of the 18S RNA subunit (Cabrera *et al.*, 2009). However, PCR does require specialised equipment and qualified personnel, limiting its use in field diagnosis.

1.2.3 Trypanotolerance

Trypanotolerance is defined as the genetically determined ability of some breeds of cattle to limit the multiplication of the parasite and to resist the pathogenic effects of infection (Murray *et al.*, 2004). Some *Bos taurus* cattle [including N'Dama (longhorn), Mutura, Baoule (short horn) and Dahomey breeds], are naturally resistant to trypanosomosis and are termed trypanotolerant (Authié, 1994). Trypanotolerance in these breeds of cattle, as with resistance in wild animals, is thought to be due to co-evolution with African trypanosomes. The humpless trypanotolerant *Bos taurus* cattle, migrated to tropical Africa much earlier (3000 BC) than Zebu cattle (*Bos indicus,* 700 AD), which are more

common in sub-Saharan Africa but trypanosusceptible (Hanotte *et al.*, 2000; Naessens *et al.*, 2002).

Trypanotolerant animals are able to (by differing extents) reduce the levels of parasitaemia and anaemia which are the main pathogenic effects of trypanosomosis, and generally, to self-cure (Black *et al.*, 2001; Guirnalda *et al.*, 2007). On a molecular level, β_2-microglobulin was found to be significantly differentially expressed in tolerant and susceptible cattle and further studies are being conducted to assess the role of this protein in trypanotolerance (Berthier *et al.*, 2003; Courtin *et al.*, 2008).

Bos indicus cattle (Zebu and Boran breeds), which were brought to the African continent much later than the taurine breeds, differ in degrees of susceptibility to trypanosomes but typically all breeds need treatment. Only one-third of all cattle in tsetse-infested regions in Africa, and less than 10% of all sub-Saharan cattle, are trypanotolerant. Trypanotolerant N'Dama cattle are of a smaller size than *Bos indicus* breeds and, therefore, thought to be less productive. As a result, use of the larger, susceptible *Bos indicus* breeds is far more widespread (Murray *et al.*, 2004).

1.2.4 Control of trypanosomosis

1.2.4.1 Tsetse control

Four main methods have been employed in Africa to control tsetse flies: aerial and ground spraying, artificial baits, including traps and targets which involve the use of screens impregnated with insecticide (Brown, 2008; Rayaisse *et al.*, 2010), insecticide treated cattle (Bouyer *et al.*, 2009) and sterile insect technique (Vreysen, 2001). These techniques are most effective when used in conjunction with each other rather than individually or integrated with anti-parasitic methods (Aksoy, 2003). Spraying can be effective when used over a large area but tends to leave behind small areas in which the flies are able to live and proliferate. Therefore, the use of bait technologies, after spraying, allows these smaller foci to be targeted directly and eliminated (Torr *et al.*, 2005). The treatment of cattle with insecticide is the cheapest method of tsetse control but does not work well in heavily infested areas. The main problem with the use of spraying and bait - which are not continuous - is the reinvasion from an untreated area since tsetse flies are able to travel up to 5 km per day (Torr *et al.*, 2005).

Sterile insect technique requires the repeated release of sterilised male insects into the wild. Over a period of time, this procedure leads to a reduction of fertile males in the

population, hence fewer successful matings and, as this continues, the population eventually becomes extinct. This technique has been used successfully against the screwworm fly in parts of North America, the Mediterranean fruit fly in California and, most recently, *Glossinia austeni* in Zanzibar (Aksoy, 2003). The Pan-African Tsetse and Trypanosomosis Eradication Campaign (PATTEC) has put forward a project for large scale tsetse fly control using sterile insect technique and insecticide treated cattle with USD $ 80 million in funding from the African Development Bank (Torr *et al.*, 2005). Nevertheless, doubts remain about the feasibility of the project over the greater than 10 million km^2 that are infested by tsetse flies (Rogers and Randolph, 2002).

1.2.4.2 Trypanocides

Three drugs have been used for the past fifty years to control AAT, viz. isometamidium, homidium and diaminazene aceturate (Figure 1.5) (Steverding, 2010). Isometamidium is used primarily as a prophylactic and is protective for approximately six months. Homidium is usually used therapeutically but can also be used as a short term prophylactic, while diaminazene aceturate is used solely for therapeutic purposes (Anene *et al.*, 2001). Although the exact mechanisms by which these drugs work are unclear, it is known that isometamidium and diaminazene target kinetoplast DNA, while homidium interferes with glycosomal functions, trypanothione metabolism and replication of kinetoplast mini-circles (Holmes *et al.*, 2004). It has also been shown that diaminazene inhibits the parasite serine peptidase oligopeptidase B (Morty *et al.*, 1998).

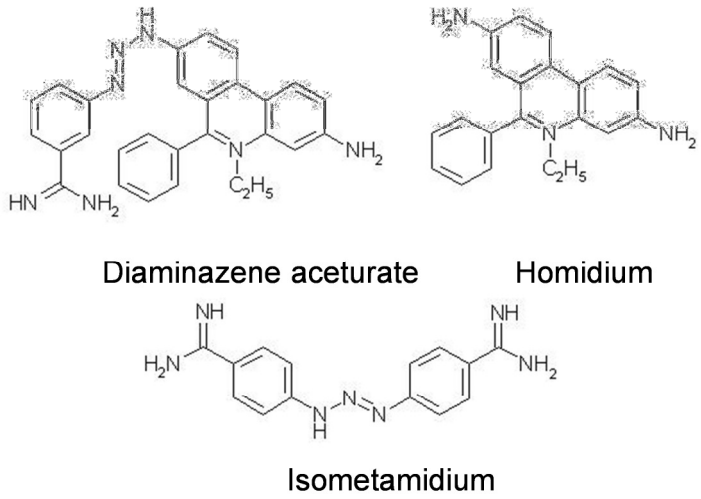


Figure 1.5 Chemical structures of the three primary drugs used to treat African animal trypanosomosis.

The development of new trypanocides is necessary for several reasons: firstly, resistance to all three of these drugs has been reported and documented in at least thirteen (one

report indicates seventeen) sub-Saharan countries (Geerts and Gryseels, 2000; Delespaux et al., 2008). Interestingly, a recent study has indicated that it may be possible to circumvent the issues of drug resistance in *T. congolense* by use of an antibiotic in conjunction with isometamidium chloride. Although this current strategy is impractical due to the cost of the large amounts of antibiotic required, it may be the first step in reversing drug resistance in trypanosomes in sub-Saharan Africa (Delespaux *et al.*, 2010).

Trypanocides are predominantly supplied by European pharmaceutical companies and are quite costly. African companies produce generics which are cheaper, but are often of unreliable quality (Holmes *et al.*, 2004). As a result, farmers are seldom convinced of the efficacy of the drugs. Another motivation for the development of new trypanocides is that the three drugs in use are all potentially toxic to humans (Finelle, 1979; Holmes *et al.*, 2004). The main obstacle in the development of new trypanocides is purely economic: the total market value for new drugs against AAT is approximately USD $ 30 million, the cost for the development of a single novel compound, however, is estimated at USD $ 250 million (Delespaux *et al.*, 2008).

Many compounds have been put forward as possible new candidates for drug design in parasitic protozoa, including glycolytic enzymes (Verlinde *et al.*, 2001), enzymes involved in trypanothione (Heby *et al.*, 2007; Spinks *et al.*, 2009; Torrie *et al.*, 2009) and purine metabolism (el Kouni, 2003), cysteine peptidases (McKerrow et al., 1999; Cazzulo et al., 2001; Alcala-Canto et al., 2007; Sharma, 2007; Ettari et al., 2010) and aminopeptidases (Knowles, 1993; Acosta et al., 2008; Gardiner et al., 2009; Skinner-Adams et al., 2009) to name a few. Note that the latter two will be further discussed in Sections 1.4 and 1.5 respectively.

1.2.4.3 Vaccines

Due to the emergence of drug resistance, vaccine development is the most viable alternative for control of AAT. There are many criteria for the development of a vaccine against parasites, including the practical considerations such as the cost per dose, biological stability, and the ease of administration (Vercruysse *et al.*, 2004). Some advances in livestock anti-parasite vaccines include vaccines for the parasites *Taenia ovis*, *Echinococcus granulosis* and *Boophilus microplus* (Dalton and Mulcahy, 2001).

Cruzain, the major cysteine peptidase of *T. cruzi,* has been shown in many studies to have high protective potential in many studies, including as a DNA vaccine (Laderach et al., 1996; Cazorla et al., 2008; Cazorla et al., 2009; Duschak and Couto, 2009). In the

case of *T. congolense*, however, immunisation of cattle with congopain did not prevent anaemia or alter the course of the infection. Yet, immunised cattle did show a recovery of PCV two to three months after infection, as well as a weight gain or maintenance of weight during infection unlike non-immunised cattle (Authié *et al.*, 2001).

Tubulin, an essential structural protein which is more than 95% identical in all pathogenic African trypanosomes, has been proposed as a vaccine candidate for these species. Refolded tubulin from *T. brucei* has shown to provide protection in mice against challenge with *T. congolense* and *T. b. rhodesiense* (Lubega *et al.*, 2002). More recently, refolded tubulin from *T. evansi* was found to provide protection in mice against *T. b. brucei, T. equiperdum* and *T. evansi* (Li *et al.*, 2007). The paraflagellar rod proteins, involved in attachment of the flagellum, have been used as a vaccine (with murine IL-12 as an adjuvant) against *T. cruzi* in mice with remarkable results (100% protection and greater than 90% reduction in parasitaemia) (Wrightsman and Manning, 2000). A third cytoskeletal protein, the microtubule associated protein 15 (MAP p15) has also been shown to have potential as a vaccine candidate. Immunisation with both native and recombinant MAP p15 form *T. b. brucei* has been shown to provide 100% protection in mice when challenged with a heterologous strain of *T. b. brucei* (Rasooly and Balaban, 2004).

The greatest obstacle to a conventional trypanosome vaccine is the phenomenon of antigenic variation, the process whereby trypanosomes alter their outer surface coat over the course of an infection (Donelson, 2003) (see Section 1.1.2.3). In an attempt to circumvent this problem, the concept of an anti-disease vaccine for parasitic diseases has been proposed (Playfair, 1991; Authié, 1994), and is discussed fully in Section 1.3.2.3.

1.3 PATHOGENIC FACTORS

Studies performed in the early 20th century showed that mice injected with heat-killed trypanosomes died. This led to the idea that certain compounds, both secreted by live trypanosomes and released by dead or dying parasites, have the ability to contribute to pathogenesis in the host (Tizard *et al.*, 1978). Moreover, infected animals show most severe symptoms of disease when the parasitaemia decreases as a result of lysis by the host immune system. This is evidence that compounds released by dead parasites contribute towards pathogenesis. Figure 1.6 shows a summary of some of the different possible pathogenic factors released by trypanosomes and the biological effects of these.

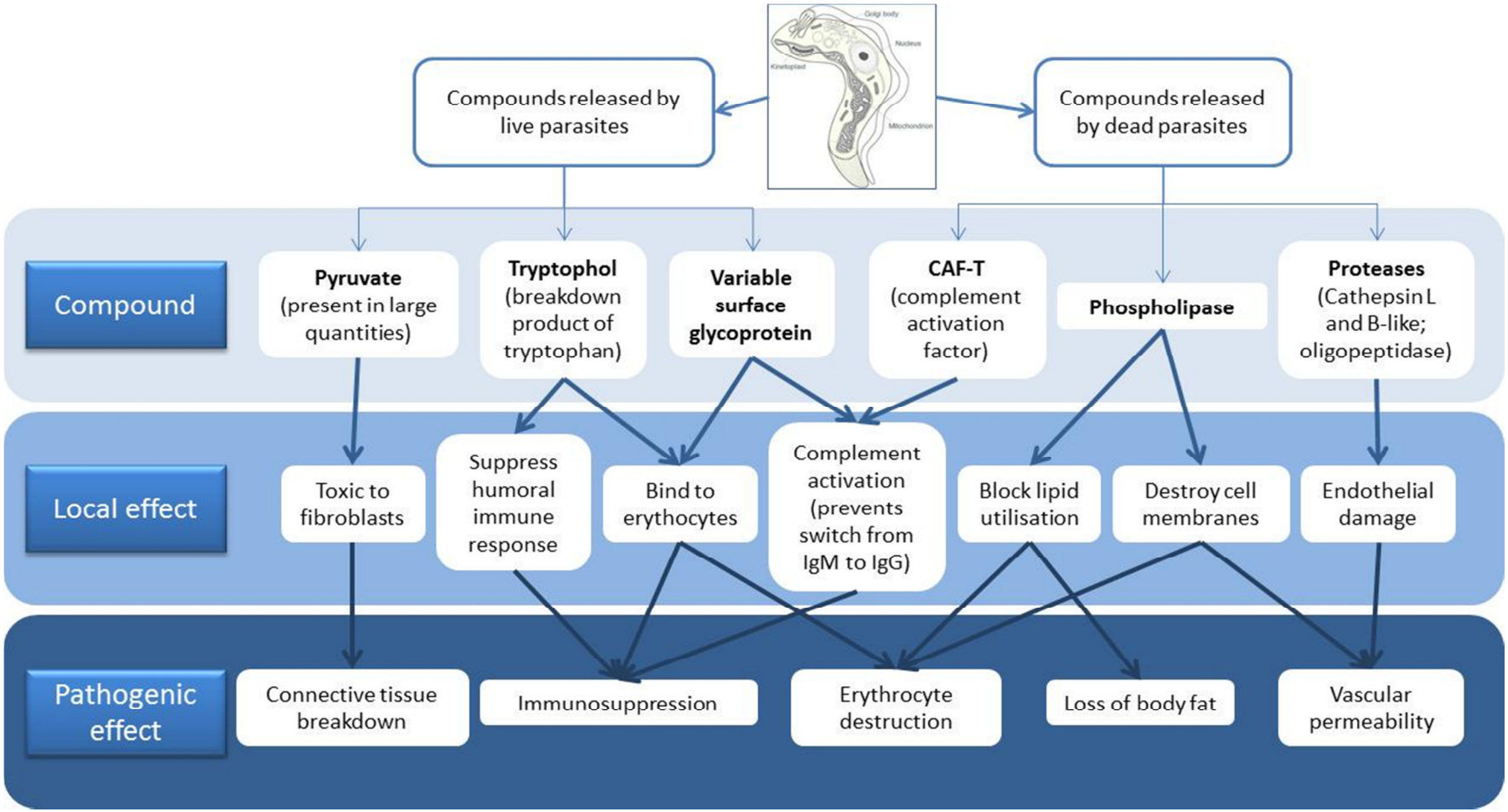


Figure 1.6 The biological roles of various compounds released by African trypanosomes which may contribute to the development of pathogenesis in the mammalian host (Tizard et al., 1978; Vincendeau et al., 1999; Wainszelbaum et al., 2001; Leppert et al., 2007; Antoine-Moussiaux et al., 2009). The effects of these compounds are further explained in the text (Section 1.3). sVSG (soluble variant surface glycoprotein), CAF-T (complement activating factor), Cat L and B-like (cathepsin L and B–like), OpdB (oligopeptidase B), RBCs (red blood cells), TNFα (tumour necrosis factor alpha), IFN γ (interferon gamma), BBB (blood brain barrier).

Also, the concept of an anti-disease vaccine (see Section 1.3.2.3) relies on the elimination of these pathogenic factors rather than the parasite itself, to moderate disease progression.

1.3.1 Compounds secreted/released from living parasites

1.3.1.1 Pyruvate

All African trypanosomes are extracellular for at least part of the period of infection in the mammalian host, and therefore, require high levels of energy for motility in the vascular system. This energy demand is supplied by an elevated level of glycolysis in bloodstream forms (Sharma *et al.*, 2009), resulting in the accumulation of large amounts of pyruvate. The excretion of this excess pyruvate may lead to connective tissue destruction in sick animals which are unable to correctly buffer the pH in their bloodstream (Tizard *et al.*, 1978).

1.3.1.2 Tryptophol

The catabolism of tryptophan by African trypanosomes via the Ehrlich pathway can produce the compounds tryptophol (indole-3-ethanol), indole-3-acetic acid and indole-2-lactic acid. It has been found that tryptophan is essential for growth and survival of the human infective trypanosomes *T. b. rhodesiense* and *T. b. gambiense*. *In vitro,* these trypanosomes take tryptophan up from the surrounding medium and *in vivo* a decrease in serum levels of tryptophan associated with human African sleeping sickness may indicate that the trypanosomes obtain their tryptophan from host serum. It has been speculated that this lowering of serum tryptophan may affect serotonin metabolism, contributing to the neurological effects of the disease (Vincendeau *et al.*, 1999).

Although the function of this pathway in trypanosomes is not known, it has been shown that injection of tryptophol into mice, rats, cats and chickens produces a coma-like state and lowered body temperature (Tizard *et al.*, 1978). Furthermore, the presence of tryptophol has been shown to reduce the humoral immune response in mice and promote the lysis of erythrocytes by binding to them rapidly and irreversibly (Tizard *et al.*, 1978). The destruction of erythrocytes is the cause of anaemia, the most debilitating symptom of trypanosomosis (see Section 1.2.1).

1.3.1.3 Soluble variant surface glycoprotein (sVSG)

Endogenous phospholipase C is capable of cleaving the GPI anchor and releasing the VSG antigen into the bloodstream of the mammalian host as soluble VSG (sVSG)

(Lamont *et al.*, 1987). Specific antibodies against the sVSG promote the formation of immune complexes in the bloodstream. The binding of these immune complexes to erythrocytes may lead to phagocytosis or lysis resulting in anaemia (Tizard *et al.*, 1978). Furthermore, it has been shown that sVSG of *T. b. brucei* activates the classical complement cascade. If this activation occurs while sVSG is bound to erythrocytes, it may also promote lysis of the erythrocytes (Tizard *et al.*, 1978). Additionally, sVSG has been shown to regulate activation of macrophages by stimulating NF-κB signalling in a murine model cell system (Leppert *et al.*, 2007). This macrophage activation also triggers the release of cytokines including TNF-α and IFN-γ, facilitating movement of the parasites across the blood-brain barrier (Antoine-Moussiaux *et al.*, 2009).

1.3.2 Compounds released from lysed parasites

1.3.2.1 Sialidases

Sialidases are enzymes responsible for catalysis of the addition or removal of sialic acid, a monosaccharide-derived group often found on surface glycoproteins. African trypanosomes may possess the ability to regulate the amount of sialic acid present on their surfaces by differential expression of a surface membrane-bound *trans*-sialidase (Montagna *et al.*, 2006). This *trans*-sialidase, expressed in procyclic form parasites, is responsible for the transfer of sialic acid groups from glycoproteins present in the tsetse fly midgut to receptor molecules, such as the side chain of GPI-anchored proteins, found on the surface of the parasite (Schenkman et al., 1994; Engstler et al., 1995). Although *trans*-sialidase activity has not been found in bloodstream forms of African parasites, there is evidence for a secreted sialidase which lacks the *trans*-sialidase activity (Nok and Balogun, 2003). The removal of sialic acid (desialyation) from macrophages may contribute to immune depression by affecting the binding capacity of the macrophages. Also, parasite sialidase-mediated changes on the surface of the host erythrocytes may result in their subsequent phagocytosis contributing to the development of anaemia in the host (Antoine-Moussiaux *et al.*, 2009).

1.3.2.2 Phospholipases

Phospholipases can be cytotoxic and haemolytic as they have the potential to rupture cell membranes by degradation of the phospholipid bi-layer. This effect is increased by the presence of free fatty acids which can be generated by the phospholipase A_1-mediated degradation of endogenous phosphatidylcholine. This process is necessary for invasion of host cells by the intracellular parasite *T. cruzi* and it has been hypothesised that the phospholipid breakdown products may be the cause of lesions found in host tissues

during infection with this trypanosome (Wainszelbaum *et al.*, 2001). The close proximity of extracellular trypanosomes to the vascular endothelium in small capillaries could therefore result in cell membrane degradation, subsequent vascular permeability and possibly, erythrocyte destruction (Tizard *et al.*, 1978). Also, it has been shown that lysophospholipids present in trypanosomes decrease the level of lysophosphatidylcholine in the mammalian host. This may contribute to weight loss by blocking the use of cholesterol and other lipids as lysophosphatidylcholine is required for both these metabolic pathways (Tizard *et al.*, 1978). Furthermore, it has been established that the non-pathogenic trypanosome species *T. theileria* has lower levels of phospholipase activity (primarily phospholipase A_1 activity) than *T. brucei* and *T. congolense,* suggesting a link between the relative levels of phospholipase activity and pathogenicity (Hambrey *et al.*, 1981).

1.3.2.3 Peptidases

The cathepsin L and B – like enzymes of several protozoan parasites have been implicated in pathogenesis, including those of *Leishmania mexicana*, *Fasciola hepatica* and *T. congolense* (Denise et al., 2003; Mottram et al., 2004; Mendoza-Palomares et al., 2008; Bryson et al., 2009; Jayaraj et al., 2009; Smooker et al., 2010). These peptidases will be discussed further in Sections 1.4.2 and 1.4.3.3. In addition, oligopeptidase B and pyroglutamyl peptidase have been associated with pathogenicity in *L. mexicana* (Munday *et al.*, 2010; Swenerton *et al.*, 2010) and *Trypanosoma* species (Morty et al., 1998; Troeberg et al., 1999; Morty et al., 2005; Morty et al., 2006; Coetzer et al., 2008).

1.3.3 Anti-disease vaccine based on pathogenic factors

An anti-disease vaccine is based on the idea that disease is caused by the pathogenic factors released by the parasite, either actively secreted or released upon death, rather than by the parasite itself. Therefore, if these pathogenic factors could be neutralised by antibodies produced by vaccination, the disease could be circumvented even if parasites are present in the host (Playfair et al., 1990; Antoine-Moussiaux et al., 2009). In order for an anti-disease vaccine to be effective, it has to be a multi-component vaccine, allowing antibody production against all factors released by the parasite which contribute to pathogenesis (Price and Kieny, 2001). Therefore, the first step in the development of an anti-disease vaccine involves the identification of pathogenic factors. Parasitic proteolytic enzymes, including the cysteine peptidases (CPs) and aminopeptidases (APs) have been implicated as pathogenic factors and are, therefore, potentially useful as components of an anti-disease vaccine (Sajid and McKerrow, 2002; Acosta et al., 2008). These two groups of peptidases will be discussed in the following sections.

1.4 CYSTEINE PEPTIDASES

Peptidases have been divided into six catalytic types based on the mechanism used in catalysis: viz. cysteine, serine, aspartate, metallo, threonine and glutamic acid peptidases (Rawlings and Barrett, 1999). The active site of peptidases contains binding sites for the substrate molecule: the site on the peptidase N-terminal to the scissile bond of the substrate is designated S_1, followed by S_2, S_3 etc., while the sites on the opposing side of the scissile bond are designated S'_2, S'_3, etc (Figure 1.7). The sites on the substrate molecule corresponding to the sites on the peptidase are designated P_1, P_2, P_3, etc. (Schechter and Berger, 1967). The side chains of the amino acid residues in P_1, P_2 etc. are designated $R_1 – R_6$.

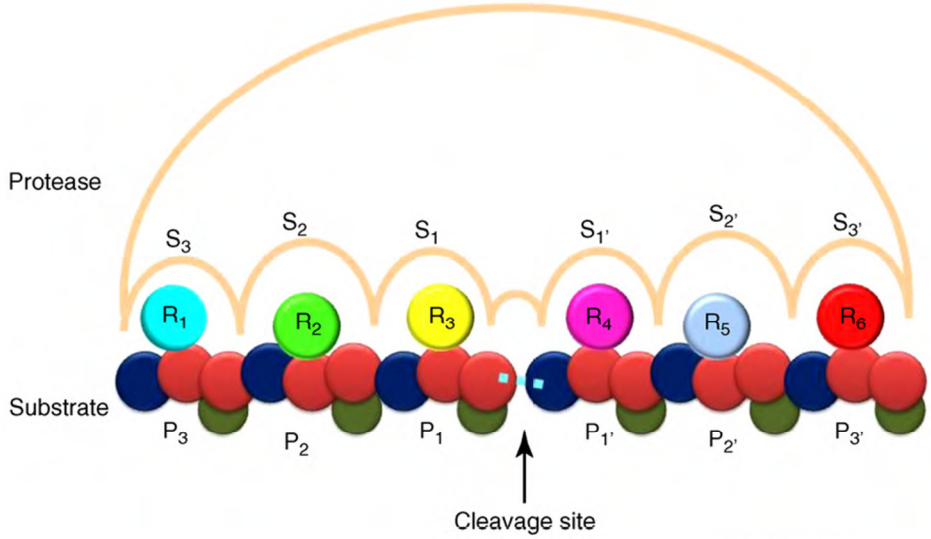


Figure 1.7 Schematic diagram of a peptidase active site annotated according to Schechter and Berger notation (Smooker et al., 2010).

1.4.1 Classification of cysteine peptidases

Cysteine peptidases (CPs) are classified into clans, which are then differentiated into families based on sequence similarity, possession of inserted peptide loops and biochemical specificity to peptide substrates (Barrett and Rawlings, 2004a; Rawlings *et al.*, 2010). Clan CA is characterised by a catalytic triad containing Cys^{25}, His^{159} and Asn^{175} (papain numbering) (Caffrey and Steverding, 2009). For catalysis by CPs, the thiol group on the catalytic cysteine residue (Cys^{25}, papain numbering) is deprotonated by an amino acid residue with a basic side-chain, usually histidine (see Figure 1.8). The anionic sulfur atom is subsequently capable of nucleophilic attack on the carbonyl group of the peptide bond forming a tetrahedral intermediate, which is stabilised by the Asn^{175} (papain numbering) residue. Following this, the tetrahedral intermediate undergoes acylation forming an enzyme-substrate thiol-ester and releasing the C-terminal portion of the substrate. The enzyme-substrate thiol-ester is hydrolysed to form a second tetrahedral

intermediate which undergoes deacylation releasing the enzyme and the N-terminal portion of the substrate (Barrett and Rawlings, 2001; Lecaille et al., 2002; Barrett and Rawlings, 2004b).

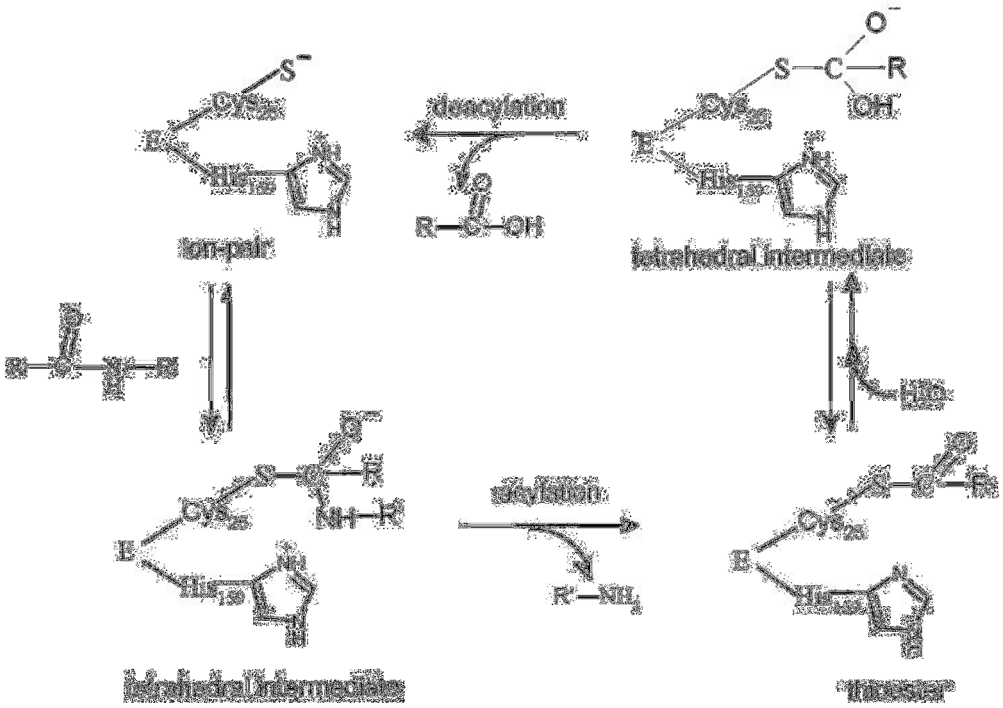


Figure 1.8 Proposed reaction mechanism for cysteine peptidases (Lecaille et al., 2002).

All Clan CA peptidases are inhibited by *trans*-epoxysuccinyl-L-leucyl-amido (4-guanidino) butane (E-64) and have a substrate specificity determined by the S_2 pocket (Barrett and Kirschke, 1981; Barrett et al., 1982). Clan CA peptidases are either targeted to intracellular vesicular compartments or secreted (Barrett and Kirschke, 1981). Family C1 of Clan CA contains the papain family CPs, that includes mammalian cathepsins L and B; and parasite peptidases, cruzain (*T. cruzi*), the falcipains (*P. falciparum*), CPA, CPB and CPC (*L. mexicana*); cathepsin L and B from *Leishmania* spp. and congopain (*T. congolense*) (Mbawa et al., 1992; Cazzulo et al., 2001; Grams et al., 2001; Barrett and Rawlings, 2004b; Zadeh-Vakili et al., 2004; Kumar et al., 2007; Beckham et al., 2009).

Family C2 of Clan CA contains the calpain-like CPs which are calcium dependent cytosolic CPs lacking a signal sequence, but possessing a calcium binding domain (Sajid and McKerrow, 2002). Note that a new unifying nomenclature incorporating the species name has been proposed for kinetoplastid papain-like CPs, but has yet to be widely adopted (Caffrey and Steverding, 2009).

1.4.2 Parasite cysteine peptidases

The CPs of parasitic protozoa typically have a higher activity at one stage of the life cycle than another; this usually being a stage in the mammalian host rather than the insect vector (Mottram et al., 1997; Lima et al., 2001; Lalmanach et al., 2002; Mundodi et al., 2002; Robinson et al., 2008; Putrianti et al., 2010). The CPs of protozoan parasites all tend to be lysosomal and are usually synthesised as preproteins. The pre-region serves as a signal peptide and is mostly hydrophobic. The pro-region is usually hydrophilic and generally conserved among members of the same species (North *et al.*, 1990).

Cysteine peptidase activity in *T. b. brucei* and *T. cruzi* has been localised in lysosome-like organelles. It has also been shown that CPs are present on the cell surface and within the flagellar pocket of *T. cruzi* (Jose Cazzulo *et al.*, 2001). Similar studies performed in *L. mexicana* have shown that CPs are present within unusual lysosome-like organelles called megasomes (Sajid and McKerrow, 2002; Caffrey and Steverding, 2009). The substrate specificity of lysosomal CPs is determined by P_2/S_2 interaction, with a preference for large aromatic groups in P_2 (Lalmanach *et al.*, 2002).

There is evidence that CPs in parasites such as *F. hepatica, Entamoeba histolytica, L. mexicana, T. cruzi* and *P. falciparum,* play a role in parasite development and differentiation, host invasion and immunosuppression, eventually contributing to pathogenesis in the host organism (Dixit et al., 2008; Robinson et al., 2008; Duschak and Couto, 2009; Quintas-Granados et al., 2009). It has been shown that a cathepsin B-like CP of *F. hepatica* (FhCB2) cleaves IgG which may contribute to immune evasion or immunosuppression (Wilson *et al.*, 1998). Also, the recombinant form of a *F. gigantica* cathepsin B-like CP (FgCB3) has been shown to degrade fibronectin, a component of the extracellular matrix. This suggests that FgCB3 may play a role in invasion of the host tissue during infection (Sethadavit *et al.*, 2009). For these reasons, CP activity-inhibiting antibodies induced by a vaccine or the use of CP inhibitors could contribute to protection against the pathogenesis associated with these parasites. However, one of the problems with designing CP inhibitors is that parasitic CPs have a similar range of substrate specificities to the mammalian host CPs (Sajid and McKerrow, 2002).

Recent evaluations of parasite vaccines comprising CPs include those of leishmaniasis, Chagas disease and fascioliasis, (Cazorla et al., 2008; Rafati et al., 2008; Jayaraj et al., 2009). Rafati *et al*. (2002) initially showed that immunisation of mice with one specific recombinant CP (rCPB) but not another (rCPA), conferred partial protection against

challenge with *L. major* in mice. In an attempt to improve the efficacy of this vaccine, Rafati *et al*. (2008) immunised mice with only the C-terminal extension of rCPA. Although the C-terminal extension proved to be extremely immunogenic in BALB/c mice, no protection was observed upon infection with *L. infantum*. Other researchers showed that immunisation with a third leishmanial CP (CPC) not only provides protection in BALB/c mice, but also significantly reduces the parasite burden (Khoshgoo *et al*., 2008). For Chagas disease, it has been shown that a DNA vaccine administered orally using *Salmonella enterica* as a delivery system provides significant protection in mice. Upon challenge with *T. cruzi* trypomastigotes, vaccinated mice showed lowered parasitaemia and reduced tissue damage (Cazorla *et al*., 2008). An initial vaccination of rats with a *F. hepatica* recombinant CP (cathepsin L1) was shown to be relatively successful, providing 78-80% protection against challenge with fluke metacercariae (Kesik *et al*., 2007). Following these promising results, immunisation and challenge of lambs and calves with the same antigen was attempted; however, only partial protection (56% at most) was achieved. Most recently, a *F. hepatica* multi-component vaccine consisting of two cathepsin L-like peptidases (cathepsins L1g and L5) and one cathepsin B-like peptidase was used for immunisation of rats. Interestingly, the highest level of protection (83%) was achieved with the omission of the cathepsin L1g peptidase (Jayaraj *et al*., 2009).

1.4.3 Cysteine peptidases of *T. congolense*

1.4.3.1 Biochemical characteristics of congopain

Cysteine peptidase activity of the metacyclic and bloodstream forms of *T. congolense* was first noted by Mbawa *et al*. (1991a) who showed that *T. congolense, T. vivax* and *T. b. brucei* exhibited differing abilities to cleave the substrate benzyloxycarbonyl-Phe-Arg-7-amino-methyl coumarin (Z-Phe-Arg-AMC) in each stage of the life cycle. Two independent studies led to the identification of congopain: Mbawa *et al*. (1991b), localised a 33 kDa CP within lysosomes which hydrolysed the synthetic substrate Z-Phe-Arg-AMC and was inhibited by cystatin and E-64. Using a cystatin affinity resin, Authié *et al*. (1992) isolated a 33 kDa antigen with the same properties. This led to the conclusion that the 33 kDa antigen and the CP were, in fact, the same molecule which came to be known as congopain (Authié, 1994). It was also established, by means of cross-reaction tests using post-infection bovine sera, that this 33 kDa CP was not present in either *T. b. brucei* or *T. vivax*. For this reason, congopain was thought to be potentially useful to diagnose *T. congolense* infections (Authié *et al*., 1992; Mbawa *et al*., 1992).

Although congopain has yet to be crystallised, the three-dimensional structure of the catalytic domain of congopain has been modelled on the crystal structure of cruzain (McGrath et al., 1995; Lecaille et al., 2001a). Congopain appears to be similar to most papain-like CPs in that it consists of two domains: an α-helical L-domain and an R-domain comprised mainly of anti-parallel β-sheets (Figure 1.9). The classical catalytic triad consisting of Cys^{25}, His^{159} and Asn^{175} (papain numbering) is located in a cleft between the two major domains (Lalmanach *et al.*, 2002). Most parasite CPs (the falcipains being a notable exception) differ from mammalian CPs by the presence of a 105 amino acid C-terminal region of indeterminate function that is linked to the catalytic domain via a polyproline (threonine in *T. vivax* and *L. major*) hinge region (Fish et al., 1995; Chagas et al., 1997). This C-terminal region has been postulated to play a role in invasion and protein targeting to a specific cellular location in some parasites (Marin-Villa *et al.*, 2008). Since the C-terminal extension is absent in mammalian CPs, it is highly immunogenic in mammals but does not induce the production of activity-inhibiting antibodies in the case of congopain (Boulangé *et al.*, 2001) or confer protection upon vaccination against *L. infantum* (Rafati *et al.*, 2008).

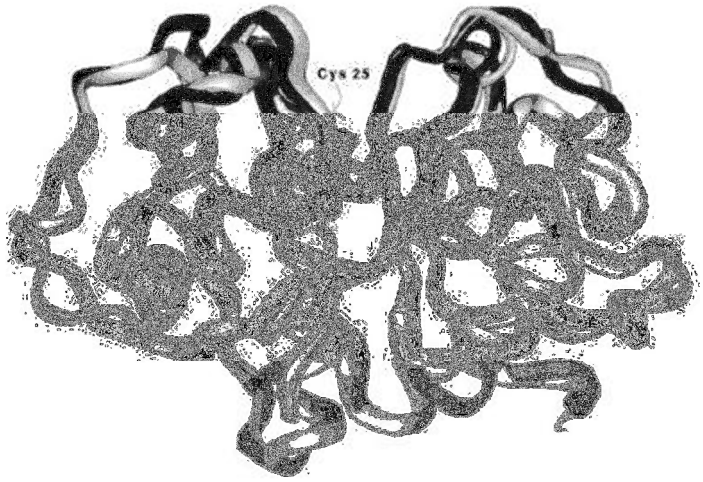


Figure 1.9 Ribbon representation of congopain modelled on cruzain, the major cysteine peptidase of *Trypanosoma cruzi* (Lecaille et al., 2001a). The active site cysteine residue (Cys25) is indicated.

Congopain, like mammalian CPs, is activated by reducing agents such as dithiothreitol (DTT), L-cysteine and β-mercaptoethanol (Mbawa *et al.*, 1992). Congopain is inhibited by low molecular weight inhibitors such as E-64 and leupeptin, and by the protein inhibitors chicken cystatin and T-kininogen (Mbawa et al., 1992; Chagas et al., 1997; Lalmanach et al., 2002). All CPs of the papain family, including congopain, prefer bulky hydrophobic residues in the P_2 position. Congopain, however, also tends to bind cyclohexylalanine better than aromatic phenylalanine derivatives (Lecaille *et al.*, 2001a). In addition, cruzain

and congopain, but not mammalian cathepsins B and L, are capable of binding a proline residue at P_2' (Serveau et al., 1996; Lecaille et al., 2001b).

Congopain, like CPs of other protozoan parasites, is located within the lysosome of the parasite and is released into the host bloodstream upon lysis of the parasite. There is a possibility that congopain may also be actively secreted (Authié *et al.*, 1993b) as it has also been localised within the flagellar pocket of *T. congolense* (Mbawa *et al.*, 1991b). Procyclic and epimastigote forms of the parasite show slight CP activity (Downey and Donelson, 1999), whereas bloodstream and metacyclic forms possess substantial amounts of cysteine peptidase activity, indicating that the CPs are developmentally regulated (Mbawa et al., 1991a; Fish et al., 1995).

Like their mammalian counterparts, trypanosome CPs have both pre- and pro-regions preceding the mature enzyme (Fish *et al.*, 1995). The pro-form of congopain can be converted into the active form of the enzyme through autolytic cleavage of the pro-region at acidic pH (pH < 5.0) (Serveau *et al.*, 2003). Autocatalytic cleavage is influenced by ionic strength and the presence of polyamines (Lalmanach *et al.*, 2002). The pro-peptide region of mammalian CPs plays a role in the proper folding of the protein, intracellular trafficking and control of the enzyme activity by blocking the active site (Lalmanach et al., 1998; Sajid and McKerrow, 2002; Caffrey and Steverding, 2009). It has recently been shown, however, that, cryptopain-1, a major CP of the parasite *Cryptosporidium parvum*, does not require the pro-domain for folding (Na *et al.*, 2009). Additionally, the pro-region of congopain, containing the inhibitory "YHNGA" peptide, has been shown not to bind to mammalian cathepsins B and L but specifically inhibit cruzain and congopain activity (Lalmanach *et al.*, 1998). In a second study Godat *et al.* (2005) showed that the level of inhibition of congopain by this peptide could be improved by using a chimeric peptide derived in part from the pro-peptide of congopain and partly from the alpha helical domain (H3) of penetratin, a transcription factor from *Drosophila antennapedia.* This discovery could possibly contribute to the design of CP inhibitory drugs (Lalmanach et al., 1998; Godat et al., 2005).

1.4.3.2 Congopain and trypanotolerance

In an experiment aimed at comparing the immune response of cattle breeds of differing susceptibility to trypanosomosis, Authié *et al.* (1993a) showed that trypanotolerant N'Dama cattle mounted a significantly greater IgG response to congopain when challenged with a primary infection of *T. congolense* compared to trypanosusceptible Boran cattle. Additional tests on F1 crosses between N'Dama and Boran cattle further

indicated that there is a correlation between the level of anti-congopain IgG produced and the degree of susceptibility to trypanosomosis (Authié *et al.*, 1993a). Hence, recognition of congopain by the immune system of certain breeds of cattle, followed by antibody production, may account for the phenomenon of trypanotolerance. If this is indeed the case, prophylaxis by means of a vaccine that includes congopain, or selective breeding using congopain as a marker for selection, could possibly lead to increased resistance to trypanosomosis in genetically susceptible cattle (Authié *et al.*, 1993a,b).

1.4.3.3 Congopain as a pathogenic factor

In order to determine whether congopain contributes to the pathogenicity of the parasite, Authié *et al*. (2001) immunised susceptible Boran cattle with two isoforms of congopain (see Section 1.4.3.4), challenged the cattle with *T. congolense,* and examined the effect of immunisation on anaemia and antibody response. By measuring the packed cell volume (PCV) or haematocrit, it was determined that immunisation did not inhibit the development of acute anaemia, but immunised animals tended to recover and attain normal PCV values two to three months after infection. The data generated suggested that congopain is likely to be a pathogenic factor since it seems to be necessary for the development of anaemia and immunosuppression in the host. However, congopain alone is insufficient as a vaccine since immunised animals developed infection and anaemia in the same manner as non-immunised animals upon challenge with *T. congolense*. Therefore, there is a need to identify other pathogenic factors which, together with congopain, may be used in an effective multi-component vaccine (Authié *et al.*, 2001).

Further evidence that congopain is a pathogenic factor is the fact that active congopain in the host bloodstream has the ability to reduce the proliferation of bovine peripheral mononuclear cells in response to mitogens and antigens (Taylor, 1998). However, other studies have found that congopain degrades various proteins *in vitro* but is unlikely to do so *in vivo,* due to endogenous peptidase inhibitors such as α_2-macroglobulin and kininogens (Lalmanach *et al.*, 2002). Conversely, Lonsdale-Eccles *et al.* (1995) have shown that kininogens actually enhance the activity of the CPs of *Leishmania* spp. and *T. b. brucei* to a small extent by lowering the pH optimum of the enzymes. In addition, a recent study has shown that the immunisation of rabbits with congopain using α_2-macroglobulin as an adjuvant increased the amount of CP-activity-inhibiting antibodies produced (Huson *et al.*, 2009). Since α_2-macroglobulin forms a complex with congopain, it is thought that the enhanced antibody production is achieved by increasing the transport of congopain to antigen presenting cells via the α_2-macroglobulin receptor (Huson *et al.*, 2009).

1.4.3.4 "Congopain-like" cysteine peptidases

There is a precedent for the cathepsin L-like CPs of parasitic protozoa to be arranged in multigenic arrays, most notably, cruzain, the major CP of *T. cruzi* (Lima et al., 2001; Duschak and Couto, 2009) and the *lmpb* gene array of *L. mexicana* and *L. donovani* which encodes structurally and functionally diverse cathepsin L-like CPs (Mottram et al., 1997; Ramos et al., 2002; Hide and Banuls, 2008). For both parasites, expression of CP isoforms differs across strains and life cycle stages. The isoforms themselves may differ in terms of glycosylation, ionic properties and substrate specificity (Klemba and Goldberg, 2002; Sijwali et al., 2006; Caffrey and Steverding, 2009).

Congopain, also known as CP2, is one of two described CPs of *T. congolense*, the other enzyme being CP1 (Fish *et al.*, 1995). While congopain has been extensively studied and characterised, native CP1 has never been isolated from *T. congolense* (Lalmanach *et al.*, 2002). The catalytic domains of the two enzymes share less identity (86%) than the pre- and pro-regions which are identical, as well as the C-terminal region (90% identity). The substitution of four basic amino acid residues for acidic ones in congopain results in a significant difference in the pI of the two enzymes: 8.35 for CP1 and 4.74 for congopain. The pH range in which the enzymes are active is also affected by this difference since CP1 is active only at acidic pH while congopain is active over a much wider pH range (Boulangé *et al.*, 2001). These biochemical differences between isoforms suggest that the sequence variability between CPs may lead to differences in antigenicity, substrate preferences and, most importantly for drug design, inhibitor specificities.

Since congopain has been shown to be a pathogenic factor (Authié *et al.*, 2001) and possibly contributes to the phenomenon of trypanotolerance, it is a definite candidate to be included in a multi-component anti-disease vaccine (Authié *et al.*, 1993a). However, this capacity for variability within the CPs of *T. congolense* has possible implications for its inclusion in such a vaccine. If multiple forms of congopain are indeed present, the various enzymes would also need to be taken into consideration for an anti-disease vaccine, or even for the production of synthetic inhibitors.

1.5 AMINOPEPTIDASES

Aminopeptidases (APs) are exopeptidases which catalyse the removal of amino acids from the unblocked N-terminus of a protein or peptide substrate. APs are ubiquitous enzymes, found in diverse organisms including plant, animal and bacterial species (Rawlings and Barrett, 1993; Mucha et al., 2010). Some of the proposed essential

functions for APs include protein maturation (Moerschell *et al.*, 1990), degradation of peptides (Botbol and Scornik, 1991; Taylor, 1993; Menzies et al., 2010), antigen trimming prior to MHC I presentation (in mammals) (Hattori and Tsujimoto, 2004; Georgiadou et al., 2010) and regulation of signal transduction (Santos *et al.*, 2000).

1.5.1 Classification of aminopeptidases

Aminopeptidases (APs) belong to the Clan MA which includes metallopeptidases such as endopeptidases, oligopeptidases and the matrix metallopeptidases that bind divalent cations (Figure 1.10). Within Clan MA, APs are found in Families M1 (alanyl APs) and M17 (leucine APs) (Rawlings *et al.*, 2010). Aminopeptidases all possess the conserved exopeptidase motif "GXMEN" which distinguishes them from the closely related leukotriene A4 hydrolases, which are also found in Family M1 (Barrett and Rawlings, 2004a). The most pertinent difference between the two families is the conserved metallopeptidase motif "HEXXH", absent in Family M17 APs, but present in Family M1 APs (Mucha *et al.*, 2010). The Glu of the motif and a Tyr residue are thought to be necessary for catalysis (Hooper, 1994). Also, APs belonging to Family M1, bind a single divalent cation, while M17 APs bind two. M17 APs are only active as homo-hexamers whereas Family M1 APs are active as individual units (Matsui *et al.*, 2006). Sub-classification of APs is based primarily on the number of amino acids cleaved from the N-terminus of substrates. Aminopeptidases cleave a single amino acid from the N-terminus while diaminopeptidases or aminodipeptidases remove a dipeptide from the N-terminus (Rawlings and Barrett, 1993).

The classification of APs is also dependent on the relative efficiency by which the enzyme removes different amino acids from the N-terminus. To date, leucine aminopeptidases (LAPs), belonging primarily to Family M17, have been prominently discussed, but arginine-, methionyl-, aspartyl-, alanyl- (mostly Family M1), glutamyl-, prolyl-, and cystinyl-APs have also been described (Barrett and Rawlings, 2004a). However, nomenclature such as "leucine aminopeptidase" may be misleading since the specificity of the enzyme *in vitro* is usually determined by the use of single amino acid synthetic substrates (e.g. H-Leu-AMC) which differ from more physiologically relevant substrates, such as peptide hormones including angiotensins (Taylor, 1993; Matsui et al., 2006; Rawlings et al., 2010). Also, although the metal ion used in activity assays appears to affect substrate preferences, different protocols use a variety of divalent cations, all of which have different effects on the catalytic activity of APs (Taylor, 1993).

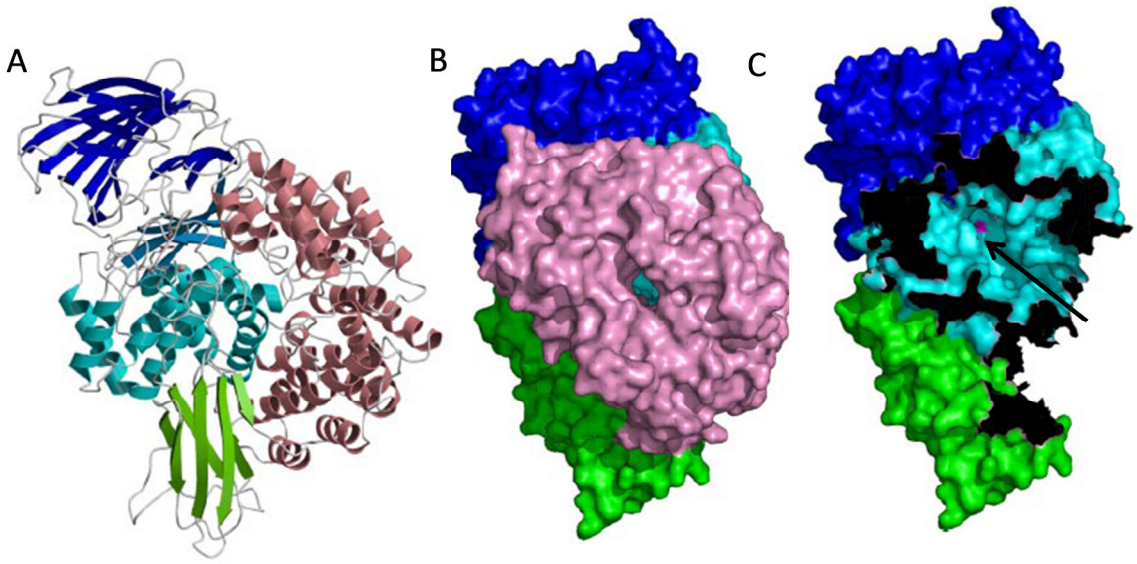


Figure 1.10 Structure of *E. coli* aminopeptidase N. (A) Ribbon model, (B) whole protein surface model and (C) protein surface model without C-terminal domain. The N-terminal domain (blue), the catalytic domain (cyan), the middle domain (green) and the C-terminal domain (pink), are shown. (C) The area of contact between the C-terminal domain and the other domains is shown in black. A magenta spot (indicated by arrow) indicates the position of the zinc cation in the active site (Taken from Ito et al., 2006).

APs may also be classified based on the location of the enzyme within the cell. Some APs are secreted, but most tend to be either membrane bound (usually Family M1) or cytosolic enzymes (usually Family M17). Also APs may be localised to a single organelle, such as lysosomes, nuclei, or mitochondria (Barrett and Rawlings, 2004a). To further confuse matters, membrane bound APs and methionine specific APs are both often referred to as aminopeptidase M. Therefore, it is possible for structurally unrelated enzymes to have the same name and be located in different parts of the cell (Rawlings and Barrett, 1993; Rawlings *et al.*, 2010). Further distinctions between APs may be made on the basis of their relative inhibition by bestatin or puromycin; metal ion content and pH of optimal activity (Taylor, 1993). APs are inhibited by bestatin, amastatin and metal-chelating agents such as EDTA and 1,10 phenanthroline (Mucha *et al.*, 2010). Puromycin specifically inhibits alanyl APs and has been used to distinguish between cytosolic alanyl APs and membrane bound APs which both belong to Family M1 (Gros *et al.*, 1985).

1.5.2 Family M1 reaction mechanism

Of the Family M1 APs, the best characterised is the *E. coli* aminopeptidase N (McCaman and Villarejo, 1982; Ito et al., 2006). The crystal structure of this AP has been elucidated (Figure 1.10) and a reaction mechanism (Figure 1.11) for the exopeptidase activity deduced (Ito *et al.*, 2006).

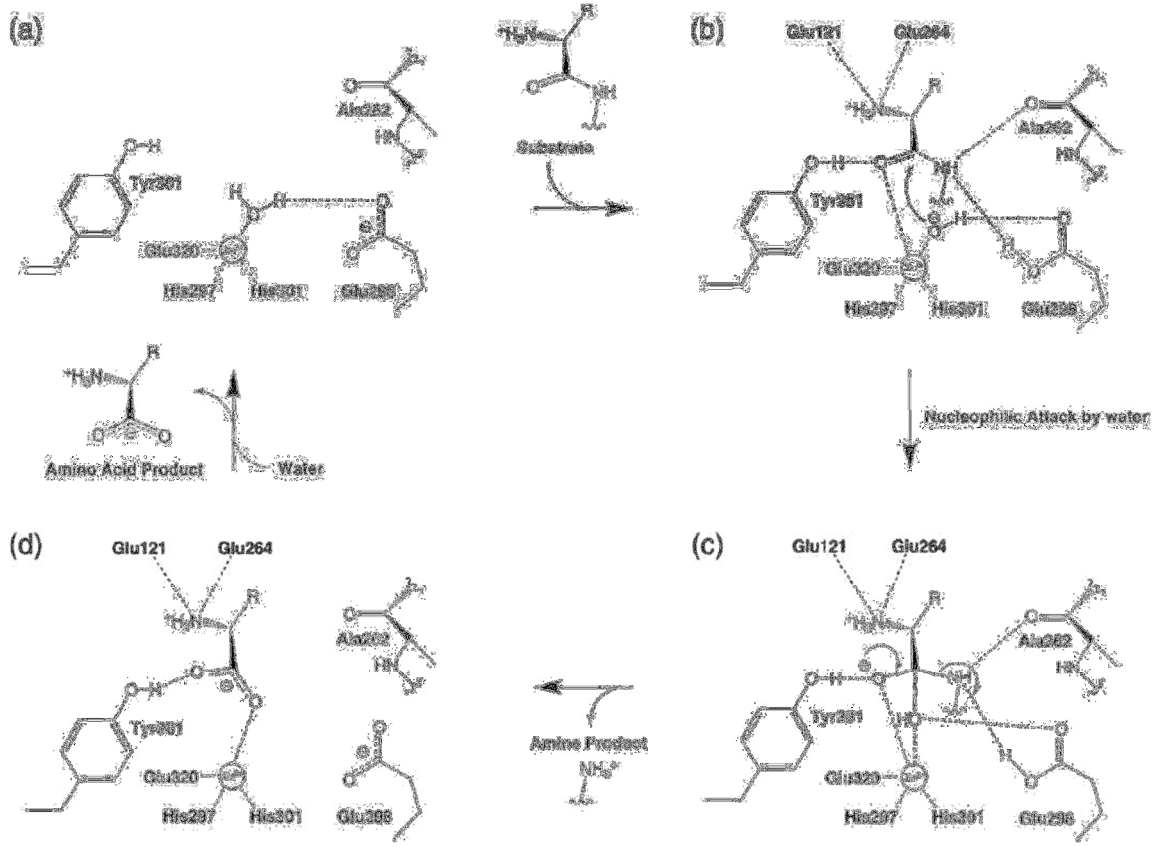


Figure 1.11 Proposed reaction mechanism for *E. coli* aminopeptidase N. It can be assumed that all Family M1 aminopeptidases have a similar mechanism for catalysis (Taken from Ito et al., 2006). (a) Free enzyme; (b) Michaelis complex; (c) tetrahedral intermediate; (d) enzyme complexed with amino acid product.

As mentioned previously, the residues necessary for catalysis are the Glu^{320} of the "HEXXH" motif and an additional Tyr^{381}, while the function of the two histidine residues, His^{297} and His^{301}, is to bind the divalent cation (Zn^{2+}). As depicted in Figure 1.11 (a), the first step of catalysis is the hydrogen bonding of Glu^{298} to a water molecule. The effect of this bond, and the proximity of the zinc ion, makes the water molecule strongly nucleophilic. The substrate is brought into the active site forming a Michaelis complex (Figure 1.11 b) by the interaction of the carbonyl oxygen of the substrate with the hydroxyl group of Tyr^{381} and Zn^{2+}. At this point, the amine group of the substrate interacts with the hydroxyl group of Glu^{298}, allowing formation of the enzyme-substrate complex.

Hydrolysis is initiated by the nucleophilic attack of the water molecule on the carbonyl carbon of the substrate (Figure 1.11 c). The hydroxyl group of Tyr^{381} is thought to stabilise this reaction intermediate. Following the nucleophilic attack by the water molecule, the amine group linked to the rest of the peptide substrate is released, leaving the N-terminal amino acid residue linked to the enzyme via the Zn^{2+} (Figure 1.11 d). This amino acid

product is released when it is replaced by a second water molecule (Ito *et al.*, 2006; Mucha *et al.*, 2010).

1.5.3 Parasite aminopeptidases

Various parasite APs have been characterised, including the APs of the protozoan parasites *P. falciparum* (Florent *et al.*, 1998), *L. amazonensis* (Morty and Morehead, 2002), *Toxoplasma gondii* (Jia *et al.*, 2010); the parasitic flukes *Schistosoma mansoni* (McCarthy *et al.*, 2004), and *F. hepatica* (Acosta *et al.*, 2007); the nematodes *Heterodera glycines* (Lilley *et al.*, 2005), *Paragonimus westermani* (Song *et al.*, 2008), *Haemonchus contortus* (Smith et al., 1997; Reszka et al., 2007) and *Setaria cervi* (Pokharel and Rathaur, 2008). Most of these studies focused on the LAPs (Family M17), except for those describing *H. glycines* and *Plasmodium* alanyl APs (Family M1). A summary of the biochemical characteristics of the recently investigated parasite APs is shown in Table 1.2.

The main role proposed for these aminopeptidases in *P. falciparum, S. mansoni* and *P. westermani* is the proteolysis of ingested haemoglobin into single amino acids, which are subsequently used for necessary metabolic processes of the parasite (Stack et al., 2007; Song et al., 2008). Given that these enzymes seem to play an essential role in parasite development, AP inhibitors seem to be effective in decreasing the growth of various intracellular parasites *in vitro*, most notably *Leishmania* spp., *T. b. brucei* and *P. falciparum* (Knowles, 1993; Morty and Morehead, 2002; Gardiner et al., 2006; Flipo et al., 2007). The most promising of these studies has produced a compound based on a malonic hydroxamic template which has proven to be a selective inhibitor of Pfa-M1, the Family M1 AP of *P. falciparum.* This inhibitor has been shown to reduce growth of parasites *in vitro* and has good physiochemical and pharmo-kinetic properties (Flipo *et al.*, 2007; Skinner-Adams *et al.*, 2009).

There have been several attempts to use APs as vaccines against parasitic diseases, most notably for fascioliasis and haemonchosis (Piacenza et al., 1999; Acosta et al., 2007; Reszka et al., 2007). Interestingly, the fascioliasis vaccination included two cathepsin L-like CPs together with the LAP of *F. hepatica*. The group receiving the mix of all three enzymes was significantly protected (78%), while the best level of protection (89%) was achieved by vaccination with the LAP alone (Piacenza *et al.*, 1999). In addition, the LAP of *F. hepatica* has been shown to be an immuno-dominant antigen and could potentially be used for the serological diagnosis of fascioliasis in humans (Marcilla *et al.*, 2008).

Table 1.2 Summary of the biochemical properties of recently characterised parasite aminopeptidases. (N/A – information not available).

Species	Family	Regulation of expression	Localisation	Substrate specificities	Inhibitor specificities	Effect of divalent cations	pH optima	References
Fasciola hepatica	M17	All stages	Cytosolic	H-Leu-AMC +++ H-Met-AMC + H-Cys-AMC+	1,10 phenanthroline DTT, PMSF, pepstatin, bestatin	Mg^{2+}, Mn^{2+}, Co^{2+} ↑↑ Zn^{2+} ↑ Ni^{2+}, Fe^{2+}, Cu^{2+} no effect Ca^{2+} ↓	8.0-8.5	(Acosta *et al.*, 2007)
Haemonchus contortus	M1	N/A	Transmembrane	L-Glu-pNa L-Leu-pNa	1,10 phenanthroline bestatin, amastatin	N/A	N/A	(Smith et al., 1997; Reszka et al., 2007)
Heterodera glycines	M1	All stages	Gut associated	N/A	N/A	N/A	N/A	(Lilley *et al.*, 2005)
Leishmania amazonensis	M17	All stages	Cytosolic	H-Leu-AMC +++ H-Met-AMC + H-Cys-AMC+	bestatin, apstatin, actinonin, arphamenine A, leucinol, 1,10 phenanthroline	Mn^{2+}, Ni^{2+}, Co^{2+} ↑↑ Zn^{2+} no effect Fe^{2+},Ca^{2+}, Cu^{2+} ↓	8.5	(Morty and Morehead, 2002)
Paragonimus westermani	M17	N/A	Gut associated	H-Leu-AMC +++ H-Arg-AMC+ H-Tyr-AMC+	1,10 phenanthroline EDTA, bestatin	Mn^{2+}, Mg^{2+} ↑↑ Zn^{2+} ↓	8.0	(Song *et al.*, 2008)
Plasmodium falciparum	M17	All erythrocytic stages, esp. trophozoite	Cytosolic	H-Leu-AMC +++ H-Phe-AMC +	1,10 phenanthroline, bestatin	Mn^{2+},Co^{2+} ↑↑ Ni^{2+}, Mg^{2+}, Zn^{2+} ↑ Cu^{2+} ↓	8.5	(Stack *et al.*, 2007)
Plasmodium falciparum	M1	All erythrocytic stages, esp. trophozoite	Cytosolic	H-Leu-AMC	N/A	N/A	7.4	(Florent et al., 1998; Allary et al., 2002)
Schistosoma mansoni	M17	N/A	Gut associated	H-Leu-AMC	1,10 phenanthroline EDTA, bestatin	Mn^{2+}, Mg^{2+} ↑↑ Zn^{2+} ↓	8.25	(McCarthy *et al.*, 2004)
Setaria cervi	M17	All stages	N/A	L-Leu-pNa +++ L-Glu-pNa+	1,10 phenanthroline bestatin, amastatin	Mg^{2+}, Mn^{2+}, Co^{2+} ↑↑ Ni^{2+},Fe^{2+},Cd^{2+}, Cu^{2+} ↓	9.0	(Pokharel and Rathaur, 2008)
Toxoplasma gondii	M17	N/A	Cytosolic	H-Leu-AMC +++ H-Ala-AMC ++ H-Arg-AMC+	bestatin	Mn^{2+}, Co^{2+} ↑↑ Zn^{2+} ↑ $Mg^{2+,}$ Ni^{2+}, Fe^{2+}, Cu^{2+} no effect	8.0	(Jia *et al.*, 2010)

1.6 OUTLINE OF THE STUDY

The overall aim of this study was to identify and characterise novel pathogenic factors of *T. congolense* for ultimate use in an anti-disease vaccine. The initial requirement, therefore, was to identify novel antigens in *T. congolense* which could contribute to the development of disease in the host. The method developed here for identification of novel pathogenic factors relies on the concept that trypanotolerant cattle are able to mitigate the disease by generating a specific immune response against a few key antigens (pathogenic factors), whereas susceptible cattle lack this ability (Authié, 1994; Authié *et al.*, 2001).

In this study the specific antigens recognised by sera from *T. congolense*-infected trypanotolerant (N'Dama) and trypanosusceptible (Boran) cattle were compared by two proteomic techniques viz. two dimensional electrophoresis and blotting, and an immuno-affinity column (Chapter 2). In the case of the two dimensional electrophoresis, *T. congolense* lysate was separated on large two dimensional gels and transferred to a PVDF membrane. Sera from *T. congolense*-infected N'Dama and Boran cattle were meant to be used as primary antibodies and the differences in profiles compared. However, due to several technical difficulties with the method, as well as the limited amounts of *T. congolense*-infected N'Dama and Boran sera available, it was not feasible to use this method to identify specific antibody recognised antigens.

The second method involved the construction of two affinity columns: one containing pre-immune (naive) sera and a second column containing sera from trypanotolerant N'Dama cattle serially infected with *T. congolense*. *T. congolense* lysate was first passed over the affinity column containing the naive sera and the flow-through, depleted of non-specific antigens, passed over the affinity column containing the sera from infected cattle. The eluate of both columns was analysed by SDS-PAGE and mass-spectrometry (MS) to identify specifically recognised antigens i.e. those found in the eluate of the column containing the sera from infected cattle but not in the eluate from the column containing the naive sera. The genes coding for the proteins identified by MS were identified using the GeneDB database [http://www.genedb.org/genedb/tcongolense (accessed 5-12-2008)].

The most promising genes identified included those coding for the cathepsin L-like CPs (multiple variants) and Family M1 Clan MA APs (2 genes). Due to the known variability of the cathepsin L-like CPs, a study of the genetic organisation was conducted (Chapter 3). Using several techniques including PCR and database mining coupled with sequencing of CP genes allowed identification of several interesting catalytic residue mutations as well as providing an in depth study of the variability of the cathepsin L-like CPs in *T. congolense*.

Two particular CP variants were cloned and expressed in the eukaryotic expression system, *Pichia pastoris* (Chapter 4). The recombinant enzymes were purified and proved to retain their functionality. The variant CPs were subsequently characterised biochemically, using the similarly recombinantly expressed catalytic domain of congopain (C2) as a reference. Papers published from this work are included at the back of the thesis.

Chapter 5 describes the expression, purification and characterisation of two APs of *T. congolense* which were identified in Chapter 2 as being potential pathogenic factors. The APs were expressed as insoluble inclusion bodies in bacterial and as a result, had to be refolded and purified. Once the refolding of the APs was validated using several techniques including an activity assay, the two recombinant enzymes were characterised biochemically. In an attempt to elucidate the *in vitro* function of the APs, the expression of the two genes was down-regulated in procyclic parasites using RNAi. Part of this work was conducted in the Laboratory of Prof. Théo Baltz (Laboratoire de Microbiologie Cellulaire et Moléculaire et Pathogénicité, UMR-5234 CNRS, Bordeaux, France).

The findings of this study are summarised and discussed in Chapter 6 (General Discussion).

CHAPTER 2: IDENTIFICATION OF NOVEL PATHOGENIC FACTORS OF *TRYPANOSOMA CONGOLENSE* USING PROTEOMIC TOOLS

D. Pillay[1*], A. Boulangé[2], T. Baltz[3] and T.H.T. Coezter[1]

[1] School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal (Pietermaritzburg campus), Private Bag X01, Scottsville 3209, South Africa
[2] Trypanosomoses, UMR-IRD/CIRAD 017, CIRAD-Emvt, 34398 Montpellier, France
[3] Laboratoire de Microbiologie Cellulaire et Moléculaire et Pathogénicité, UMR-5234 CNRS, Bordeaux, France
[*] Email: 202511575@ukzn.ac.za

Trypanosoma congolense **is a protozoan parasite which causes African animal trypanosomosis (AAT), a wasting disease of cattle and small ruminants. Current control methods are inadequate resulting in severe economic losses due to AAT in sub-Saharan Africa. The ability of African trypanosomes to alter their surface coating during the course of an infection (antigenic variation) greatly hinders the development of a conventional vaccine. An anti-disease vaccine approach against trypanosome infection has been proposed, whereby the parasite pathogenic factors responsible for the development of disease in the host are targeted, rather than the parasite itself. For this approach to be successful, it requires the identification of these pathogenic factors. In the present study, IgG from *T. congolense* infected trypanotolerant cattle was isolated and coupled to an affinity matrix. Passing a *T. congolense* lysate depleted of non-specific antigens over this affinity column allowed identification of antigens specifically recognised by the immune system of infected trypanotolerant cattle. Following this approach, immunoprecipitation of *T. congolense* antigens has allowed the identification of several potential pathogenic factors, including cathepsin L - like enzymes and two M1 type aminopeptidases.**

Trypanosoma congolense is a protozoan parasite that infects cattle and small ruminants causing African animal trypanosomosis (AAT). The parasite is transmitted by the tsetse fly that infests most of sub-Saharan Africa. AAT is of great concern, resulting in economic losses from USD $ 1 - 5 billion per annum in Africa (Kristjanson *et al.*, 1999). Development of AAT in the host mammal usually begins with the rapid proliferation of parasites under the skin in the area surrounding the tsetse bite, sometimes resulting in a chancre. The parasites enter the lymphatic system and subsequently the bloodstream causing severe anaemia and immunosuppression, leaving the infected animal susceptible to other opportunistic infections. Without treatment, infection with *T. congolense* can result in a chronic infection or death, depending on the strain of *T. congolense* and the genetics of the host animal (Taylor and Authié, 2004).

There are currently three main methods of controlling AAT: the use of trypanocides, controlling the tsetse fly populations, and the use of livestock which are innately resistant to the disease (trypanotolerant) (d'Ieteren et al., 1998; Anene et al., 2001; Gooding and Krafsur, 2005). Three trypanocidal drugs *viz.* isometamidium, diaminazene aceturate and

homidium have been available for the past fifty years. However, these drugs are costly, and drug resistant parasites have been found in 17 African countries (Delespaux *et al.*, 2008). The control of tsetse fly populations by destruction of their habitat and the use of insecticides has been effective but is not ecologically sustainable (Torr *et al.*, 2005). Sterile insect technique has been successfully employed to eradicate smaller island populations of tsetse flies (Vreysen *et al.*, 2000) but is unlikely to be effective over the 10 million km^2 infested with tsetse fly in sub-Saharan Africa (Rogers and Randolph, 2002). Trypanotolerant livestock have an innate genetic ability to limit the pathogenic effects of trypanosome infection, and, as a result, are able to survive in areas infested by tsetse (Naessens *et al.*, 2002). However, due to their smaller size and perceived lower production, farmers are unwilling to use trypanotolerant cattle, preferring the larger susceptible breeds (Murray *et al.*, 2004).

The phenomenon of antigenic variation, the mechanism through which the parasite is able to alter its outer surface protein and evade the immune system, has been an obstacle to the development of a conventional vaccine against AAT. For this reason, an alternate vaccine strategy, called an anti-disease vaccine, has been proposed for AAT (Authié, 1994; Authié et al., 2001; Lalmanach et al., 2002; Antoine-Moussiaux et al., 2009). An anti-disease vaccine targets the pathogenic factors responsible for the development of disease, rather than the parasite itself (Tizard *et al.*, 1978). However, for an anti-disease vaccine to be successful, all compounds of the parasite which contribute to pathogenesis would need to targeted (Playfair *et al.*, 1990; Schofield, 2007). Therefore, the first step in the development of an anti-disease vaccine is the identification and characterisation of these pathogenic factors.

As mentioned, some breeds of cattle (notably N'Dama) possess an innate genetic resistance to trypanosomosis known as trypanotolerance (Naessens et al., 2002; Murray et al., 2004). It has previously been shown that *T. congolense*-infected N'Dama cattle produce antibodies against some antigens which the immune system of Boran cattle (a susceptible breed) do not recognise (Authié *et al.*, 1993a; Authié *et al.*, 1993b). It is possible that the parasite antigens specifically targeted by the immune system of trypanotolerant animals are pathogenic factors that play a role in the development of disease in the mammalian host. Therefore, by producing a specific immune response to these critical antigens, trypanotolerant animals are able to minimise the pathogenesis of trypanosome infection. Consequently, identification of antigens specifically recognised by trypanotolerant cattle would be essential for the development of an anti-disease vaccine for AAT.

In the present study two methods were used to identify potential pathogenic factors from *T. congolense*. Firstly, western blots of two dimensional separations of trypanosome lysate were probed with sera from *T. congolense* infected trypanotolerant N'Dama or trypanosusceptible Boran cattle. However, due to several technical problems, it was not possible to identify any antigens specifically recognised by the sera of the trypanotolerant N'Dama cattle using this technique. For the second method used to identify potential pathogenic factors, IgG from *T. congolense*-infected trypanotolerant N'Dama cattle was coupled to an affinity matrix. As a control to minimise non-specific binding, IgG from a non-infected cow was similarly coupled to an affinity matrix. A lysate of *T. congolense* bloodstream form parasites was first passed over the non-immune IgG column to deplete the lysate of non-specific antigens. This non-specific antigen depleted lysate was subsequently passed over the affinity column containing IgG from *T. congolense* infected cattle. The eluates from the two columns were compared by one and two dimensional electrophoresis and the different bands and spots identified by mass spectrometry. This allowed identification of those parasite antigens specifically recognised by the immune system of trypanotolerant cattle, i.e. pathogenic factors.

2.1 MATERIALS

Trypanosomes: BALB/c mice were obtained from the National Health Laboratory Services (Pretoria, South Africa). DE-52 resin was obtained from Whatman (Middlesex, UK). Cyclophosphamide was obtained from Sigma (Munich, Germany). Protocols using mice were approved by the University of KwaZulu-Natal animal ethics committee (Reference 036/09/Animal). *T. congolense* (strain IL1180) was originally isolated from a lion in the Serengeti, Tanzania (Geigy and Kauffmann, 1973). *T. congolense* (strain IL3000) is a rodent-passaged clone of the ILC49 strain, originally isolated from a cow from the Transmara, Kenya (Wellde *et al.*, 1974). Both strains were kindly provided by the International Livestock Research Institute (ILRI), Nairobi, Kenya.

Two dimensional electrophoresis: Complete Protease Inhibitor Cocktail Tablets® were from Roche (Mannheim, Germany). Protein Assay Reagent from Bio-Rad (Hercules, CA, USA) and filter paper (N^{o} 1) from Whatman (Middlesex, UK). All reagents and equipment used for two dimensional gel electrophoresis including the solubilisation solution, 2D Clean-up Kit®, immobilised pH gradient (IPG) buffer, 3 – [(3 - cholamidopropyl) dimethylammonio] - 1-propanesulfonate (CHAPS), DeStreak Reagent®, Immobiline® Dry Strip (pH 3 - 10, non-linear, 24 cm), IPGPhor® running apparatus, enhanced

chemoluminescence (ECL) substrate kit, Typhoon 2400 scanner and the Ettan® DALT electrophoresis system were from GE Healthcare (Buckinghamshire, UK), formerly Amersham Pharmacia Biotech. Tank buffer (10x) was purchased from Euromedex (Strasbourg, France).

Immunoprecipitation: Infected sera and some *T. congolense* (strain IL3000) bloodstream form lysate was a gift from E. Authié (experiments performed at ILRI, Kenya), CNBr activated Sepharose 4 Fast Flow® and Protein G Sepharose 4 Fast Flow® were obtained from GE Healthcare (Buckinghamshire, UK).

2.2 METHODS

2.2.1 Growth and isolation of trypanosomes

A frozen stabilate of bloodstream form *T. congolense* (strain IL3000) was thawed and diluted with phosphate-saline-glucose (PSG: 57 mM Na_2HPO_4, 3 mM NaH_2PO_4, 42 mM NaCl, 50 mM glucose, 1 mM hypoxanthine, pH 8.0). Approximately 10^4 parasites were injected i.p. into either BALB/c mice or cyclophosphamide-treated (50 mg/kg 24 hours prior to infection) outbred mice. Parasitaemia was monitored daily by microscopic analysis of tail blood until the first peak of parasitaemia was reached (usually 7-10 days post infection). Mice were anaesthetised and bled by cardiac puncture and the blood-containing heparinised tubes kept on ice.

Trypanosomes were isolated from blood as described (Lanham and Godfrey, 1970). Briefly, DE-52 resin was weighed and stirred gently with a 4-fold excess of phosphate buffered saline (PBS: 100 mM Na_2HPO_4, 2 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 8.0). The slurry was allowed to settle (30 min) and the fine particles decanted. This process was repeated three times. A column was packed (approximately 4 ml resin for 1 ml of blood) and equilibrated with five column volumes of PSG. The heparinised blood was added to the top of the column with the tap closed and allowed to settle for 10 min. The tap was opened and PSG added to the top of the column to prevent the resin from drying. The flow through was checked to confirm the presence of parasites. The eluted parasites were quantified using a haemocytometer and centrifuged (3000 *g*, 10 min, 4°C) prior to storage at -80°C.

2.2.2 Two dimensional gel electrophoresis and blotting

Two dimensional gel electrophoresis allows the separation of proteins based on two parameters: molecular weight and iso-electric point (O'Farrell, 1975). Iso-electric focusing

(IEF) constitutes the first dimension of separation, and discontinuous SDS-PAGE (Laemmli, 1970), the second. *T. congolense* (strain IL3000) bloodstream form lysate (100 µg/strip) was prepared for loading onto the first dimension iso-electric focusing (IEF) strips. Protein was precipitated using the 2D Clean-up Kit® and the final pellet resuspended in solubilisation solution [2 M thiourea, 7 M urea, 4% (v/v) CHAPS, 0.1% (v/v) Triton X-100, 1% (v/v) IPG buffer, 12 µl/ml DeStreak Reagent®]. The IEF strips (Dry Strip pH 3 – 10, non-linear, 24 cm) were rehydrated with 460 µl of solubilised trypanosome lysate proteins overnight at 25°C. IEF was conducted using the protocol outlined in Table 2.1. If not used immediately, the rehydrated IEF strips were frozen at -20˚C.

Table 2.1 Protocol for the IEF step for two dimensional gel electrophoresis. Current was limited at 50 mA and the procedure conducted at 20°C.

	Voltage	Time (h)
Step 1	60 V constant	3
Step 2	60 – 1000 V gradient	4
Step 3	1000 – 8000 V gradient	4
Step 4	8000 V constant	8
Hold	40 V constant	11

In preparation for the second dimension electrophoresis, the protein-containing IEF strips were incubated with an equilibration buffer [50 mM Tris-Cl buffer, pH 8.0, 6 M urea, 2% (w/v) SDS, 65 mM DTT, 0.0002% (w/v) bromophenol blue] for 15 min with agitation to reduce the proteins and give them an overall negative charge. Addition of a second equilibration buffer, containing 260 mM iodoacetamide instead of DTT, allowed alkylation of the reduced proteins.

Polyacrylamide gels (10%, 24 x 21 cm, Ettan DALT electrophoresis system®) used for the second dimension separation, based on molecular weight, were prepared. The prepared IEF strips were rinsed in deionised water and placed on top of the second dimension gel. An agarose solution containing 0.0002% (w/v) bromophenol blue was used to seal the strips into place. Electrophoresis was conducted overnight (4°C, 20 mA/gel). A ready-to-use buffer system, TG-SDS 10x (Euromedex), was used at 2x concentration in the upper tank (2 L) and 1x concentration in the lower tank (7.5 L).

Many variables were tested concerning the transfer of separated proteins from the large gels to different membranes. Most importantly, two different buffer systems i.e.

continuous (Towbin *et al.*, 1992) and discontinuous (Laurière, 1993) were used in a semi-dry blotter using either nitrocellulose or PVDF membranes. Prior to blotting, the gel, membrane and filter papers were soaked in the transfer buffer. When using PVDF, the membrane was pre-wetted in methanol before incubation in the buffer solution.

For the continuous buffer system, a Tris-glycine buffer [20 mM Tris-Cl buffer, pH 8.8, 192 mM glycine, 20% (v/v) methanol, 1% (w/v) SDS] was used for soaking all components of the gel sandwich. Transfer was conducted at 400 mA for 90 min. For the discontinuous buffer system, four different buffers were used. Two sheets of Whatman filter paper placed closest to the cathode were incubated in buffer IV [100 mM Tris-Cl buffer, pH 10.4, 20% (v/v) methanol], followed by two sheets incubated in buffer III [20 mM Tris-Cl buffer, pH 3.8, 60 mM lactic acid, 20% (v/v) methanol], the membrane, gel, a further three sheets soaked in buffer II [25 mM Tris-Cl buffer, pH 8.4, 15 mM lactic acid, 0.1% (w/v) SDS] and finally, two sheets closest to the anode soaked in buffer I [100 mM Tris-Cl buffer, pH 8.4, 60 mM lactic acid, 0.4% (w/v) SDS]. Transfer was conducted at 400 mA for 75 min.

Following transfer, the membranes were blocked in blocking buffer [3 or 5% (w/v) non-fat milk in PBS] for 1 h, prior to incubation overnight at 4˚C with either trypanotolerant N'Dama or trypanosusceptible Boran sera diluted in blocking buffer. The membranes were subsequently washed using PBS (6 x 5 min) and incubated in secondary antibody (goat anti-bovine IgG-HRPO conjugate), also diluted in blocking buffer. Following a second washing step (PBS, 6 x 5 min), the membrane was developed using the ECL substrate kit. After removing excess substrate, the membrane was scanned at a wavelength of 428 nm in a Typhoon 2400 Scanner (GE Healthcare).

Quality of transfer was assessed by staining the membrane with Ponceau S [0.1% (w/v) Ponceau S in 1% (v/v) glacial acetic acid] and staining the gel after transfer with Coomassie Blue [0.125% (w/v) Coomassie Blue R-250, 50% (v/v) methanol, 10% (v/v) acetic acid]. The best conditions for transfer were found to be the continuous buffer system and nitrocellulose membrane. For detection, a serum dilution of 1:200 and a 1:1000 dilution of the goat anti-bovine IgG HRPO conjugate proved to be optimal. Also, it was found that by cutting the membrane into quarters or halves, more efficient incubation with primary sera could be achieved.

2.2.3 Affinity purification of IgG

Serum from naive cattle (naive IgG fraction N'Dama, 13.12.04) was previously precipitated with ammonium sulfate. This immunoglobulin (Ig) fraction was first heated at 65°C and cooled, twice, and subsequently centrifuged (12 000 *g*, 15 min) to remove any precipitate. The clarified Ig fraction (5 ml) was dialysed against 20 mM Na-phosphate buffer (pH 7.0) overnight at 4°C. Protein G Sepharose 4 Fast Flow® (1.5 ml) was washed with 5 column volumes of 20 mM Na-phosphate buffer (pH 7.0). The dialysed Ig fraction was applied to the column and washed with 5 column volumes of 20 mM Na-phosphate buffer (pH 7.0). Eight fractions (0.5 ml each) were eluted with 100 mM glycine-Cl buffer, pH 2.5. The fractions were collected into 50 µl of 1 M Tris-Cl buffer (pH 9.0) for neutralisation. The column was washed with 10 column volumes of 20 mM Na-phosphate buffer (pH 7.0) and the process repeated with the remaining Ig fraction (3.5 ml). Fractions were assayed for protein using the Bradford Protein Assay Reagent (Bio-Rad). The four fractions with the highest protein content were pooled (Pool 1) and the remaining four fractions pooled separately (Pool 2).

2.2.4 Immuno-affinity isolation of *T. congolense* antigens

IgG purified from sera from two different serially infected N'Dama cattle [ND12 (1.11.04) and ND13 (16.12.04)] was used. IgG (40 mg) was dialysed against PBS overnight at 4°C, and the buffer changed and dialysis continued for a further 2 h. The samples were centrifuged (12 000 *g*, 5 min) to remove any precipitate. The dialysed sample (40 mg) was diluted to 4 ml with PBS.

CNBr activated Sepharose 4 Fast Flow® (1 g) resin was resuspended in 4 ml of 1 mM ice-cold HCl, and washed with a further 50 ml of 1 mM HCl. The resin was split into two tubes and mixed with 4 ml (40 mg) of the prepared IgG either from naive or infected cattle and incubated at RT for 2 h with gentle agitation. The resin was centrifuged (400 *g*, 1 min, 4°C) and the supernatant discarded. The resin was washed with 10 ml PBS and incubated with 10 ml of 0.1 M Tris-Cl buffer (pH 8.0) for 2 h at RT with gentle agitation. The resin was centrifuged (400 *g*, 1 min, 4°C) and washed with 3 volumes of Na-acetate buffer (100 mM acetate, pH 3.5, 0.5 M NaCl). The washes with Tris-Cl and acetate buffers were repeated twice.

T. congolense (strain IL3000) lysate (4.5×10^9 parasites) was centrifuged (12 000 *g*, 5 min) to remove debris, and peptidase inhibitors added (PI cocktail, Roche). The lysate was mixed with the resin containing IgG from naive cattle and rotated gently overnight at

4°C. The resin was centrifuged (400 *g*, 1 min, 4°C) to collect the supernatant which was subsequently mixed with the resin containing IgG from infected cattle for 6 h at 4°C. The resin containing naive IgG was resuspended in phosphate buffer (150 mM phosphate, 0.5 M NaCl, pH 7.4) and washed with 10 volumes of the same phosphate buffer. Bound protein was eluted with 100 mM glycine-Cl buffer, pH 2.5. The eluate was dialysed against distilled H_2O overnight at 4°C. Following the incubation, the resin containing IgG from infected cattle was eluted in a similar manner.

The eluates were concentrated by lyophilisation and run on a 10% reducing SDS-PAGE gel (Laemmli, 1970) and two separate two dimensional gels as described in Section 2.2.2. For silver staining, gels were immersed in fixing solution [50% (v/v) methanol, 12% (v/v) acetic acid, 0.019% (v/v) formaldehyde] overnight and washed (3 × 5 min) in 50% (v/v) methanol. Following 1 min incubation in pre-treatment solution (4 mg/ml $Na_2S_2O_3.5H_2O$), the gel was rinsed (3 × 5 min) in distilled water and soaked in impregnation solution [0.2% (w/v) $AgNO_3$, 0.028% (v/v) formaldehyde] for 20 min. The gel was rinsed in distilled water (2 × 20 s) and incubated in developing solution [0.019% (v/v) formaldehyde, 0.004% (w/v) $Na_2S_2O_3.5H_2O$, 60 g/l Na_2CO_3] until the first bands were visible. Development was allowed to proceed in distilled water until the gel was immersed in stopping solution [50% (v/v) methanol, 12% (v/v) acetic acid].

The bands/spots of interest were excised from the gel and subjected to trypsin digestion. The resultant peptides were analysed by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) for peptide mass fingerprinting which allowed the sequence of the individual peptides to be identified by comparison to the *T. congolense* genome sequence hosted at GenBank® (http://www.ncbi.nlm.nih.gov/genbank/, accessed 4-12-2008). (Note that excision of bands/spots from the gel, trypsin digestion and LC-MS/MS was conducted by the Plateforme Génomique Fonctionnelle Bordeaux, France).

2.3 RESULTS

The first method used in an attempt to identity novel pathogenic factors from *T. congolense* involved two dimensional gel separation and blotting (Figure 2.1). Several major problems were encountered with this technique. Firstly, blotting large gels (24 x 20 cm) was found to be not very efficient and varying degrees of protein transfer was achieved with each blot, as could be seen by Ponceau S staining (results not shown). Still, it was essential to have equal quantities of protein on each blot since not only the identity, but also the relative quantities of the proteins recognised differentially by sera

from trypanotolerant N'Dama and trypanosusceptible Boran cattle were to be compared. The second problem was linked to the scarcity of suitable sera. Given that the optimal dilution of sera was determined to be 1:200, and that at least 10 ml of this dilution was necessary for each large membrane, it was not possible to repeat these experiments several times for optimisation. In addition to this, the sera available was over ten years old (1994) and had been subjected to a few freeze-thaw cycles by this time. For these reasons, no conclusive results of differentially recognised antigens could be gained from this method.

Figure 2.1 shows that the VSG proteins of approximately 50 kDa (indicated by solid line) are recognised by both Boran (Figure 2.1 B) and N'Dama (Figure 2.1 C) sera. Furthermore, it is evident that the sera from infected trypanosusceptible Boran cattle seemed to recognise more *T. congolense* antigens than the sera from infected trypanotolerant N'Dama cattle.

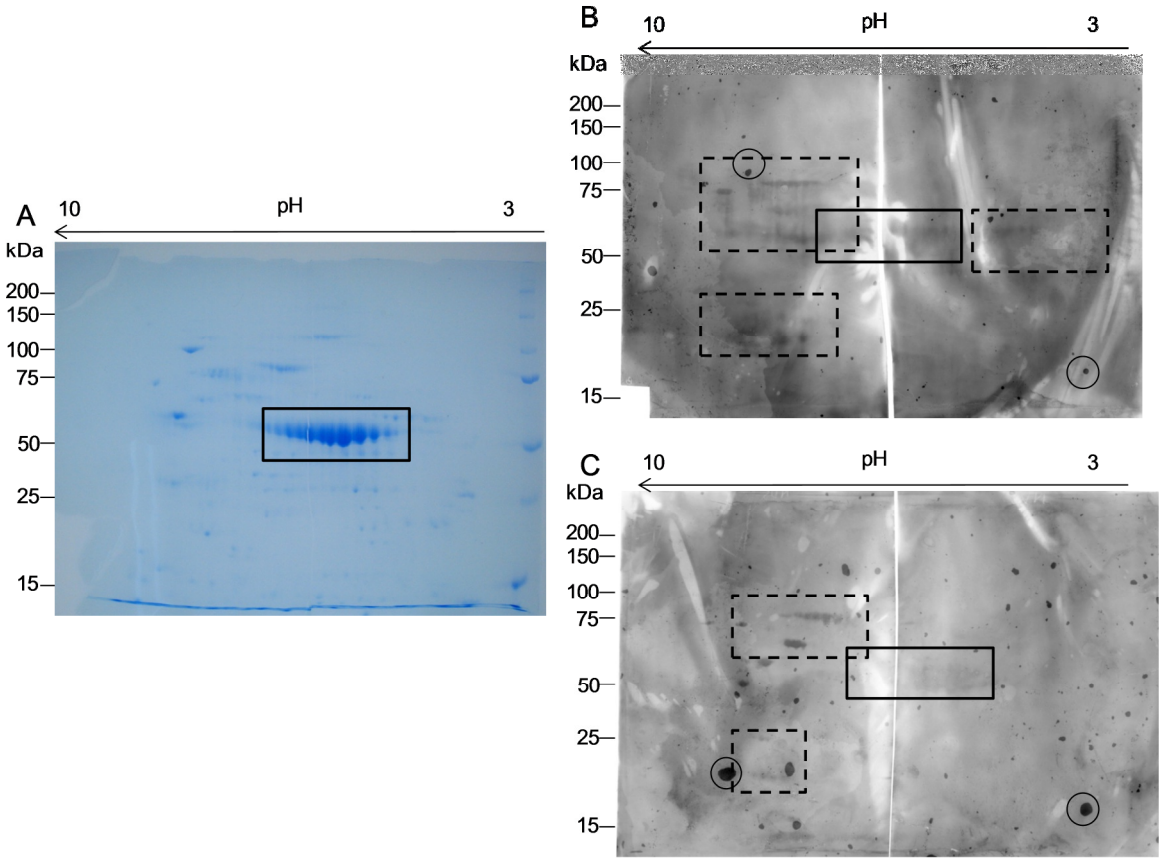


Figure 2.1 Two dimensional separation of *T. congolense* lysate. (A) Gel stained with Coomassie blue as reference. Western blot of two dimensional separation of *T. congolense* lysate probed with (B) infected trypanosusceptible Boran serum (1:200) and (C) infected trypanotolerant N'Dama serum (1:200). Goat anti-bovine IgG-HRPO conjugate (1:1000) was used as a secondary antibody. Smooth lines indicate the VSG and dashed lines indicate other recognised (but unidentified) proteins. Circles indicate the presence of air bubbles. Arrows indicated pH gradient.

For the second approach used to identify potential pathogenic factors of *T. congolense,* two affinity columns, containing IgG from either naive or *T. congolense*-infected cattle were prepared. Trypanosome lysate was passed over each of these immuno-affinity columns in turn and the eluates from each compared by reducing SDS-PAGE (Figure 2.2). The gels were silver-stained in order to visualise low quantities of proteins. Four distinct bands at approximately 100, 45, 30 and 25 kDa were of different intensities in the eluates from the two immuno-affinity columns containing either naive or infected IgGs. For the 100 and 45 kDa bands, substantially more material was present in the eluate from the column containing IgG from infected cattle. It is likely that the band between 50 and 60 kDa, represents the VSG protein. The bands were excised from the gel as indicated on Figure 2.2 and analysed by LC-MS/MS by the Plateforme Génomique Fonctionnelle Bordeaux, France.

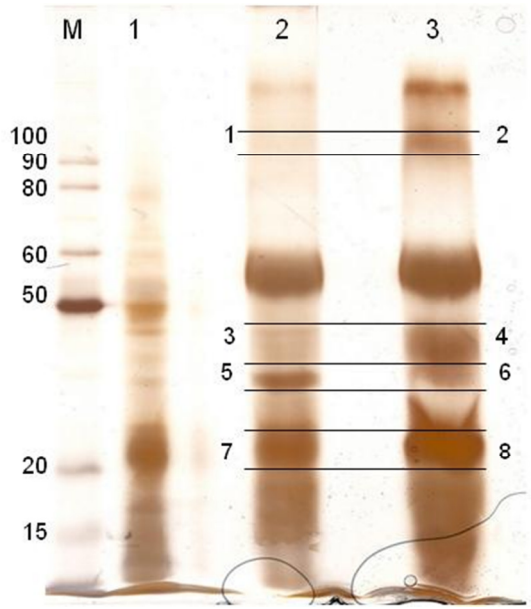


Figure 2.2 Silver stained 10% reducing SDS-PAGE gel showing the results of the trypanosome lysate elutions from columns containing IgG from either naive or *T. congolense* infected cattle. Lane M: Invitrogen Benchmark® molecular weight marker (5 µl); lane 1, total trypanosome lysate (5 µg); lane 2, eluate from column containing IgG from naive cow column (5 µg); lane 3, eluate from column containing IgG from infected cattle (5 µg). Bands of differing intensity in the two eluates are indicated.

LC-MS/MS analysis allowed identification of the major proteins found only in the eluate from the column containing the infected IgG but not in the eluate of the column containing naive IgG (Table 2.2). The “peptides covered” column in Table 2.2 refers to the number of trypsinised peptides from each particular band that corresponds to the identified antigen. It can, therefore, be concluded that these proteins were specifically recognised by IgG from the *T. congolense* infected cattle. Of these proteins, the cysteine peptidases and aminopeptidases were selected for further studies.

Table 2.2 Major antigens specifically recognised by IgG from infected cattle.

Band	Major antigen(s)	Peptides covered	Molecular weight (kDa)
100 kDa	aminopeptidase putative	5/6	96.8
	aminopeptidase putative	8/12	97.3
45 kDa	protein phosphatase 2C putative	2/3	45.8
	phosphoglycerate kinase	3/3	44.9
	arginine kinase	2/2	40.0
	unknown protein	3/4	44.7
30 kDa	protein phosphatase 2C putative	2/2	40.1
	cysteine peptidase	3/4	33.4
	cysteine peptidase	2/2	48.5
	cysteine peptidase	3/4	33.6
25 kDa	cysteine peptidase	3/3	33.6

The fractions eluted from the columns containing IgG from *T. congolense* infected cattle and naive cattle were also analysed by two dimensional electrophoresis (Figure 2.3 A and B). Two sets of spots with approximate molecular weights of 30 and 35 kDa (circled on Figure 2.3 A) were present only in the eluate from the column containing IgG from infected cattle. Analysis of these spots by LC-MS/MS indicated that the 30 kDa spots were cysteine peptidases, similar to those identified using one dimensional electrophoresis (see Table 2.2). The even horizontal spacing between these spots on Figure 2.3 A indicated that these peptidases differ equally from each other by a unit of charge, since the horizontal dimension resolves proteins according to their isoelectric points. The 35 kDa spots were identified as fragments of the structural protein, tubulin.

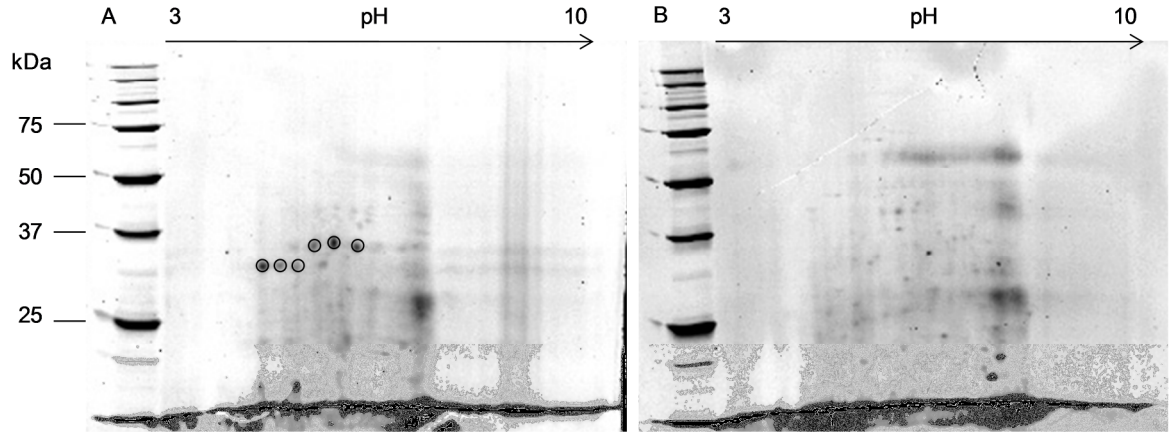


Figure 2.3 Silver-stained two dimensional reducing gel showing the results of the trypanosome lysate elutions from columns containing IgG from either *T. congolense* infected or naive cattle. Two dimensional separation of eluates from columns containing IgG from (A) *T. congolense* infected cattle and (B) naive cattle. Circle spots were unique to the eluate from the column containing IgG from *T. congolense* infected cattle but not in the eluate from the column containing IgG from naive cattle. Arrows indicate pH gradient.

2.4 DISCUSSION

African animal trypanosomosis (AAT) is a wasting disease of cattle and small ruminants which is of great concern in sub-Saharan Africa. Since the current control methods are inadequate, and the development of a conventional vaccine is greatly hindered by the phenomenon of antigenic variation, the concept of an anti-disease vaccine for AAT has been proposed (Authié et al., 2001; Antoine-Moussiaux et al., 2009). The first step in creating such a vaccine would be to identify all parasite factors involved in pathogenesis.

The two methods employed to identify potential pathogenic factors (two dimensional blotting and immuno-affinity isolation), involved determining which antigens are strongly recognised by IgG from infected trypanotolerant (N'Dama) cattle while excluding those antigens also recognised by IgG from trypanosusceptible (Boran) cattle. This is based on the rationale that one of the possible explanations behind the phenomenon of trypanotolerance is that tolerant cattle are able to mount a specific immune response against those factors of the parasite responsible for pathogenesis, specifically anaemia (Authié et al., 1993a; Authié, 1994). Concerning parasitaemia, trypanotolerant cattle are also able to modulate the amount of parasites but this does not seem to play as significant a role in the progression of the disease (Naessens, 2006).

A proteomic approach was used by Holzmuller *et al*. (2008a) to analyse trypanosome strains of differing virulence. In their study, lysates of two strains of *T. b. gambiense* which showed differing clinical symptoms in mice were separated by two dimensional electrophoresis and the protein profiles compared. Spots which were different were excised from the gel, proteolytically digested and the peptides identified using MS. These peptides were then used to identify the corresponding gene by searching the sequenced *T. brucei* genome at GeneDB. It was evident that although the genomes of the two strains are practically identical, different patterns of expression of key genes could contribute to the level of virulence of different *T. b. gambiense* strains (Holzmuller *et al*., 2008a). A second study compared *T. congolense* strains IL1180 and IL3000 and three *T. evansi* strains, Teva1, TeGub-323 and El Frío, using proteomic tools. For *T. congolense*, the majority of proteins differentially expressed between the two strains were either peptidases or glycolytic enzymes (Holzmuller et al., 2008b; Grébaut et al., 2009). A similar study which analysed the proteomes of three different strains of *T. cruzi,* also showed that cruzain, the major cysteine peptidase, was an important marker for strain virulence (Kikuchi *et al*., 2010). These results validate the proteomic technique as a means of identifying these key antigens, since both these groups of proteins (peptidases

and glycolytic enzymes) have shown promise as either drug targets or vaccine candidates in protozoan parasites (Verlinde et al., 2001; Sajid and McKerrow, 2002; Sharma, 2007; Duschak and Couto, 2009).

Most recently this method of whole proteome analysis has been used to identify 444 secreted proteins in three strains of *T. b. brucei* and to compare the differences in expression in procyclic forms of *T. b. brucei* and *T. b. gambiense* (Geiger *et al.*, 2010). These combined results have shown that trypanosomes may secrete proteins via an exosome pathway which could be useful for developing a vaccine which prevents transmission from the vector to the mammalian host (Atyame Nten *et al.*, 2010). A similar study of the *T. evansi* proteome was conducted to identify which proteins are expressed during infection, as well as identifying new antigens with diagnostic potential. Since the *T. evansi* genome has never been sequenced, it was necessary to perform homology searches of the *T. cruzi* and *T. brucei* sequence databases. Nevertheless, over 100 antigens were identified using this technique (Roy *et al.*, 2010). These powerful proteomic techniques are very useful for the identification of proteins which may be markers of virulence or drug resistance, proteins specific to each life cycle stage, or immuno-dominant antigens which are useful for diagnosis (Cuervo *et al.*, 2010).

The second method used to identify potential pathogenic factors from *T. congolense* involved immuno-affinity isolation of *T. congolense* antigens using sera from infected cattle. To this end, IgG isolated from serially infected trypanotolerant N'Dama cattle were coupled to an affinity matrix. In order to exclude those antigens which are naturally recognised, a second column was constructed using IgG isolated from naive N'Dama cattle. Whole trypanosome lysate (bloodstream forms) was passed, first over the column containing IgG from naive animals to eliminate non-specific antigens and subsequently over the column containing IgG from infected animals. Several antigens were identified as being recognised specifically by sera from infected animals, including two M1 type aminopeptidases, two putative 2C protein phosphatases, a phosphoglycerate kinase, an arginine kinase and several cysteine peptidases. The success of this fundamental "one-dimensional" method, compared to the more cumbersome and expensive two-dimensional method emphasises the value of basic techniques in identifying new virulence factors worthy of further study.

The Family M1 type aminopeptidases are metallo-exopeptidases that play a role in degradation of peptides to liberate amino acid residues for use in protein synthesis

(Taylor, 1993; Hooper, 1994; Hattori and Tsujimoto, 2004). Although Family M1 aminopeptidases have been used in a vaccine for *Fasciola hepatica* (Piacenza et al., 1999; Acosta et al., 2008) and have shown promise as a drug target in *Plasmodium falciparum*, (Chen et al., 2006; Flipo et al., 2007; McGowan et al., 2009; Skinner-Adams et al., 2009; Mucha et al., 2010) this group of peptidases has never been studied in the pathogenic trypanosomes (Family M1 aminopeptidases are discussed at length in Section 1.5).

In mammals, protein phosphatases 2C (also known as serine/threonine-specific phosphatases) are essential for processes such as glycogen metabolism, muscle contraction and protein and fatty acid synthesis (Cohen, 1989). Protein phosphatases and kinases are enzymes which catalyse the removal and addition respectively of phosphate groups from other molecules. Both these groups of enzymes have been extensively studied in trypanosomatids (Brenchley *et al.*, 2007; Szoor, 2010) and both play an essential role in cell signalling. The low level of sequence identity between parasite and mammalian kinases and phosphatases and the fact that these enzymes are essential for survival of the parasite suggest that they could possibly be useful as drug targets (Naula et al., 2005; Doerig et al., 2010).

Cysteine peptidases are proteolytic enzymes which have been widely studied in protozoan parasites and have stimulated interest as both vaccine candidates and drug targets (Sajid and McKerrow, 2002; Caffrey and Steverding, 2009; Ettari et al., 2010). When used as a vaccine against trypanosomosis in cattle, the major cysteine peptidase of *T. congolense* provided partial protection (Authié *et al.*, 2001). A complicating factor in the use of cysteine peptidases of parasitic protozoan as either drug targets or vaccine candidates is that these enzymes tend to be present as multi-copy genes which have significant sequence differences (Denise *et al.*, 2003; Dvořák *et al.*, 2005; Mendoza-Palomares *et al.*, 2008). The exact nature of the multiple copies of cathepsins L-like cysteine peptidases in *T. congolense* has not yet been explored. (Cathepsin-L like cysteine peptidases are discussed more fully in Section 1.4).

The aim of this study was to identify novel potential pathogenic factors from *T. congolense* using two different proteomic tools, viz. two dimensional gel separation and blotting; and immunoprecipitation. Several problems were encountered with the two dimensional gel separation and blotting technique, and as a result, no significant conclusions could be drawn from these experiments. The immuno-affinity isolation technique, on the other

hand, allowed several potential pathogenic factors to be identified and of these, two groups of peptidases, the aminopeptidases and the cysteine peptidases, were selected for further study.

CHAPTER 3: ANALYSIS OF THE VARIABILITY OF THE CONGOPAIN-LIKE CYSTEINE PEPTIDASES OF *TRYPANOSOMA CONGOLENSE*

D. Pillay[1*], A. Boulangé[2] and T.H.T. Coezter[1]

[1] School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal (Pietermaritzburg campus), Private Bag X01, Scottsville 3209, South Africa
[2] Trypanosomoses, UMR-IRD/CIRAD 017, CIRAD-Emvt, 34398 Montpellier, France
[*] Email: 202511575@ukzn.ac.za

Congopain is the major cathepsin L-like peptidase of *Trypanosoma congolense,* the prime causative agent of trypanosomosis in cattle in Africa. It has previously been shown that the congopain gene exists as a multigenic family consisting of 2 kb gene units organised in a tandem array. At least one genetic variant of congopain, known as CP1, has been characterised but never isolated from the parasite itself. This variability has implications for the inclusion of congopain in a multi-component vaccine. Two different mini-libraries of congopain-like genes were prepared: one which selected for genes as different as possible from congopain, and a second which contained all possible genes present in the congopain locus. Analysis of the sequences obtained in these two mini-libraries showed that there is significant variability of the genes within the congopain array. The presence of isoforms of congopain *in vivo* was confirmed by probing a two dimensional separation of the cytoplasmic proteins of *T. congolense* with monoclonal antibodies specific to two different regions of congopain. Recognition of multiple spots by these monoclonal antibodies is an indication that at least some of the nucleotide variability observed from the sequencing of the libraries is expressed as proteins.

Congopain, also known as CP2, is the major cathepsin L-like cysteine peptidase (CP) of *Trypanosoma congolense*, the causative agent of bovine trypanosomosis. Congopain has previously been identified as a pathogenic factor since the enzyme has been linked to the development of anaemia and immunosupression in the host (Authié *et al.*, 2001; Lalmanach *et al.*, 2002). In addition, it is thought that congopain may play a role in the phenomenon of trypanotolerance, the inherited genetic resistance to trypanosomosis displayed by some West African breeds of *Bos taurus* cattle (Authié, 1994). For these reasons, congopain is a likely candidate for an anti-disease vaccine or target for synthetic inhibitors. Congopain, however, like the cathepsin L-like peptidases of *Leishmania mexicana* and *T. cruzi,* appears to be present in multiple isoforms (Fish et al., 1995; Mottram et al., 1997; Lima et al., 2001; Cancela et al., 2008).

Indeed, CP1, a divergent congopain-like peptidase (Figure 3.1) has been described in *T. congolense* (Fish et al., 1995; Boulangé et al., 2001; Boulangé et al., 2011)but unlike congopain, native CP1 has not been isolated from the parasite to date. This variability has implications for the development of a vaccine or inhibitor using congopain as a target.

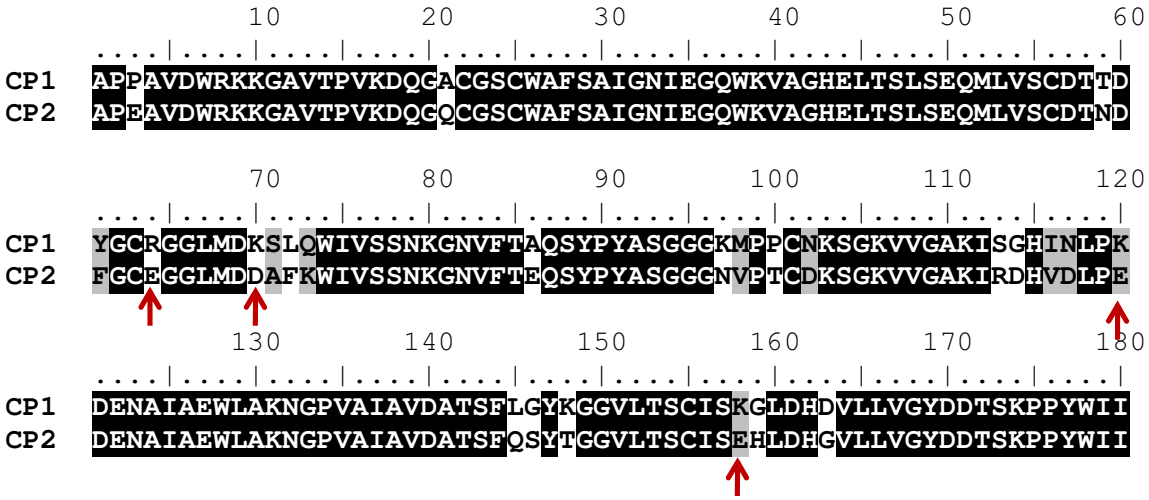


Figure 3.1 Alignment of part of the catalytic domains of CP1 and congopain (CP2). Arrows indicate substitutions of acidic amino acid residues (Glu and Asp in congopain/CP2) for basic residues (Arg and Lys in CP1); resulting in a difference in pI for the two enzymes. Sequences obtained from GenBank® (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucleotide, accessed 5-9-2007), GenBank ID Z25813 and L25130 respectively. Sequences aligned using the ClustalW subprogram (Chenna *et al*., 2003) within BioEdit (Hall, 1999).

Two dimensional gel electrophoresis is a powerful technique for the separation of heterogeneous protein containing samples, such as whole cell lysates. When this technique was first developed, 1100 *E. coli* proteins could be successfully resolved, and the system was expected to resolve up to 5000 proteins (O'Farrell, 1975). In addition to providing an estimate of the total number of different proteins present in a sample, two dimensional gel electrophoresis allows resolution of proteins differing by a single charge. As an example, the substitution of four acidic amino acid residues in congopain (CP2) for four basic residues in CP1 (Figure 2.1) results in a significant difference in the pI of the two enzymes: 4.74 for congopain and 8.35 for CP1 (Lalmanach *et al*., 2002). Trypanosome genes are polycistronic (Pays, 2005). For this reason, mRNA transcripts do not correlate directly with expressed protein, suggesting that protein expression is regulated at a translational level rather than a transcriptional level (El-Sayed *et al*., 2000). Therefore, studies of mRNA transcripts are limited in terms of validating protein expression, making two dimensional electrophoresis an essential tool for assessing gene expression (van Deursen *et al*., 2003).

In a study aimed at elucidating the genomic organisation of congopain genes in a given *T. congolense* (strain IL3000) (Wellde *et al*., 1974), Boulangé *et al*. (2010) showed by Southern blotting experiments that congopain-like CPs exist as a multigene family of 2 kb gene units organised as a tandem array (Figure 3.2). This study was followed by the construction of a cosmid library of 105 clones of 40-50 kb fragments of *T. congolense* genomic DNA. Analysis of these recombinant cosmids indicated that each gene unit is

2 kb in length, consisting of a coding region and a highly conserved intergenic region (Boulangé et al., 2011). Additionally, a cleavage site for the enzyme MluI was found to be conserved in the intergenic region of all gene units allowing whole gene units to be separated and cloned individually (see Figure 3.3).

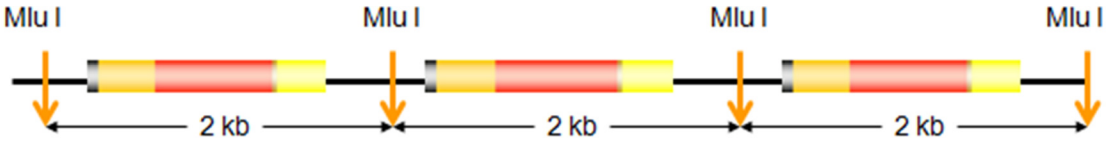


Figure 3.2 Schematic diagram of the arrangement of genes within the congopain gene array. Digestion of this locus with MluI releases 2 kb fragments, each containing an entire gene unit [Diagram courtesy of A. Boulangé, University of KwaZulu-Natal]

Due to the high degree of conservation of the intergenic region of the gene array, conventional PCR amplification of any particular gene is impossible. Therefore, a DNA library-based approach is necessary to clone all possible variants, and gain access to any selected gene. This approach entails the construction of sub-genomic libraries enriched in gene units (2 kb fragments) coding for congopain-like CPs. The DNA can be obtained from either the clones selected from the recombinant cosmid library, or by PCR using *T. congolense* genomic DNA as a template, or directly from size-digested genomic DNA. In this study, the first two of these approaches have been attempted.

This study aims to show that variants of congopain are indeed present in the parasite by use of two dimensional gel electrophoresis. Also, to analyse this variation at the gene level, firstly by studying the sequences in databases of the partially sequenced *T. congolense* genome, and secondly by creating sub-genomic DNA libraries.

3.1 MATERIALS

Two dimensional electrophoresis: Complete Protease Inhibitor Cocktail Tablets® was from Roche (Mannheim, Germany). Protein Assay Reagent was obtained from Bio-Rad (Hercules, CA, USA). Filter paper (N^{o} 1) was obtained from Whatman (Middlesex, UK). All reagents and equipment used for two dimensional gel electrophoresis including the solubilisation solution, 2D Clean-up Kit®, immobilised pH gradient (IPG) buffer, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), DeStreak Reagent®, Immobiline® Dry Strip (pH 3 - 10, non-linear, 24 cm), IPGPhor® running apparatus, enhanced chemoluminescence (ECL) substrate kit, Typhoon 2400 scanner and the Ettan® DALT electrophoresis system were from GE Healthcare (Buckinghamshire, UK), formerly Amersham Pharmacia Biotech. Tank buffer (10x) was purchased from Euromedex (Strasbourg, France). Monoclonal antibodies 4C5 (Authié *et al.*, 1992) raised against

native congopain and 3B10 (T. Lefrançois, unpublished, CIRAD, Montpellier, France), raised against the recombinantly expressed catalytic domain of congopain (C2), were obtained from A. Boulangé (University of KwaZulu-Natal, Pietermaritzburg and CIRAD). Goat anti-mouse IgG horse radish peroxidase (HRPO) conjugate was obtained from Sigma-Aldrich (Munich, Germany).

Molecular biology: The following molecular biology reagents were obtained from Fermentas (Vilnius, Lithuania): MluI, EcoRI, NotI, shrimp alkaline phosphatase (SAP), T4 DNA ligase, 10 mM dNTP mix, X-gal, MassRuler® DNA Ladder Mix; FastRuler® DNA ladder, GeneJet® Plasmid Miniprep Kit and O'RangeRuler® DNA Ladder 200bp. *Taq* polymerase, 10x PCR reaction buffer and $MgCl_2$ solutions were obtained from Solis Biodyne (Tartu, Estonia). The Wizard® PCR Preps DNA Purification System, Wizard® Plus SV Minipreps DNA Purification System, Wizard® Genomic DNA Purification Kit, and the pGEM-T® Easy cloning kit were obtained from Promega (Madison, WI, USA). pMul1, a pGEM-11Zf+® vector altered by site-directed mutagenesis to contain a cleavage site for the enzyme MluI (see Figure 3.3), was obtained as a glycerol stock from E. Mulinge (Kakundi, 2008). Recombinant cosmids containing *T. congolense* genomic DNA (strain IL3000) were provided by A. Boulangé (University of KwaZulu-Natal). Sequencing was performed at the SegoliP Sequencing Unit at the Institute for livestock research (ILRI) in Nairobi, Kenya.

***E. coli* cells:** Competent *E. coli* cells (NEB 5-alpha K12 or One Shot® Top10F') were purchased from New England Biolabs (Ipswich, MA, USA) or Invitrogen (Carlsbad, CA, USA). Both *E. coli* strains used allow blue/white screening for transformants in the presence of X-gal by α-complementation of the β-galactosidase gene.

3.2 METHODS

3.2.1 Two dimensional electrophoresis and blotting

Conducted as described previously. (See Section 2.2.2).

3.2.2 Bioinformatic methods

Sequences homologous to congopain were obtained by conducting a BLASTN (Altschul *et al.*, 1990) search of the *T. congolense* reads database hosted at GeneDB (http://www.genedb.org/genedb/tcongolense/, accessed 5-9-2007), using the nucleotide sequence of congopain (GenBank® accession: L25130). Sequences were aligned with the Sequencher® software (Gene Codes Corporation, 2006), using the congopain

nucleotide sequence as a reference sequence, the Dirty Data algorithm, a match percentage of 75%, and a minimum overlap of 20%. Degenerate primers needed for the construction of the libraries were designed based on this alignment Table 3.1. The alignment was exported to BioEdit© Sequence Alignment Editor where the nucleotide sequences were translated into amino acid sequences in the specified frame (Hall, 1999). ClustalW was used to align the amino acid sequences allowing sequence differences to be compared (Chenna *et al*., 2003). All sequences subsequently obtained from the libraries were aligned in a similar manner.

Table 3.1 Primer sequences used for constructing a CP PCR library from amplification of *T. congolense* (strain IL3000) genomic DNA and for cloning and expression of multiple *T. congolense* CPs. Primer set 1 designed based on the MluI restriction site (underlined) in the intergenic region of the congopain gene unit (Figure 3.2) used in the construction of a CP library by PCR. Primer sets 2 and 3 designed for cloning and expression of the amplicon incorporating an EcoRI site in the forward primer and a NotI site in the reverse primer (underlined). The same forward primer was used for sets 2 and 3. The T-vector primers (T7 and SP6) were used to confirm the presence of the insert in the T-vector. Congopain specific primers were used to screen the libraries for genes most different to congopain.

Set	Amplicon	Forward (5'3')	Reverse (5'3')
1	2 kb (whole gene)	5'ACCGTGTGTCTTTCTATRC3'*	5'CACGCGTGGTTGGGGCRGC3'*
2	1260 bp (full-length ORF)	5'GCCGAATTCGCGTGCTYTGTTCCCGTGGCG3'*	5'CCAGGCGGCCGCCTCGTGYCGCACGAGCCGRGC3'*
3	960 bp (catalytic domain)	5'GCCGAATTCGCGTGCTYTGTTCCCGTGGCG3'*	5'TAGCGGCCGCACCCTCGCASAACTCCTGCGTGA3'*
T7 SP6	Variable (T-vector MCS)	5'TAATACGACTCACTATAGGG3'	5'TATTTAGGTGACACTATAG3'*
CP2 specific	670 bp	5'GGAGGCATTCCGTTTCCG3'	5'CCATGGTCCAGATGCTCG3'*

* Degenerate bases labelled according to IUPAC/IUB symbols: R (A/G); Y(C/T) and S (G/C).

3.2.3 Construction of CP cosmid library from recombinant cosmids

Fifteen cosmid clones selected in a previous study using the congopain gene sequence as a probe were digested with MluI (Boulangé, unpublished). This restriction enzyme cuts in the intergenic region of the congopain gene array to release the 2 kb gene units. The restriction enzyme digest was analysed on an agarose gel [0.8% (w/v)] and the band corresponding to 2 kb excised and purified. All agarose gels were prepared as 1% (w/v) agarose, unless otherwise stated. A molecular weight marker (1 µl) of an appropriate size range was loaded in the first lane of all agarose gels; sample (1 µl) was loaded (unless otherwise stated); gels were run at 80 V in 1x TAE buffer containing ethidium bromide and

photographed under ultraviolet light. pMul vector was prepared as follows: *E. coli* transformed with the vector was streaked from a glycerol stock onto a 2x YT agar plate containing ampicillin (100 µg/ml) and grown overnight at 37˚C.

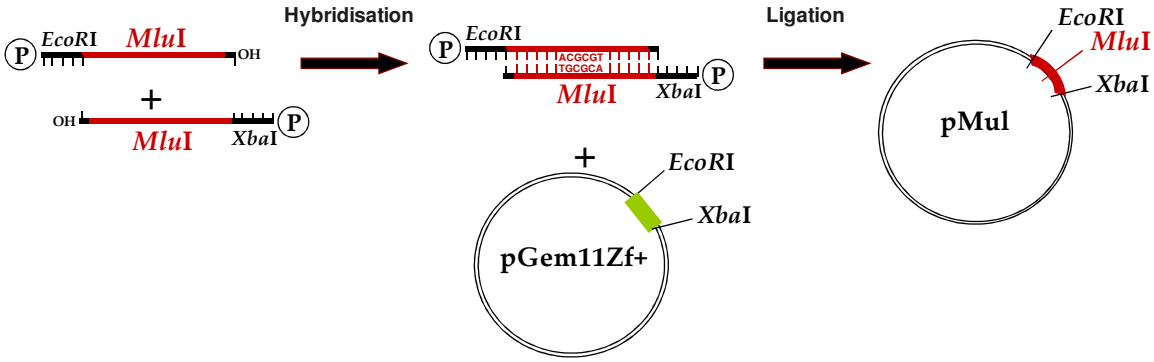


Figure 3.3 Schematic diagram showing construction of the pMul vector (Kakundi, 2008). A synthetic DNA fragment was cloned into the plasmid pGEM-11Zf+. Two phosphorylated oligonucleotides comprising a MluI site, homologous to each other except for EcoRI and XbaI overhangs, were allowed to hybridise, and the double-stranded fragment thus generated was ligated into pGEM-11zf+ previously digested with EcoRI and XbaI. Oligonucleotides were designed in such a way that the *lacZ* gene fragment is not interrupted, allowing blue/white screening in pMul.

A single colony was picked and inoculated into 2x YT liquid media containing ampicillin (100 µg/ml) and grown overnight at 37˚C. The pMul vector was isolated by miniprep and digested with MluI and subsequently dephosphorylated using SAP according to the manufacturer's instructions. The 2 kb fragment purified from the MluI digested cosmids was ligated into the linear, dephosphorylated pMul vector using T4 DNA ligase at 4˚C overnight.

An appropriate fraction of the ligation mix (no more than a tenth of the cell volume) was used to transform commercial (K12 or Top10F') competent *E. coli* cells by heat shock (42˚C, 30 s). Following an outgrowth period of exactly 1 h (to limit the possibility of mutations arising) at 37˚C with agitation, the cells were plated onto 2x YT agar plates containing IPTG (15 µl of 0.1 M per plate) and X-gal (10 µl of 50 mg/ml per plate) and grown overnight at 37˚C. White colonies, indicative of a recombinant plasmid were screened by colony PCR using the T7 and SP6 primers specific to the vector from which pMul was constructed. PCR was conducted in a 20 µl volume as per Table 3.2. The PCR (20 µl) was analysed on an agarose gel. Glycerol stocks of confirmed recombinants were created by mixing overnight liquid culture of the clone with an equal volume of sterile 50% (v/v) glycerol and stored at -70˚C. This process was repeated until a total of 44 clones were obtained. DNA sequencing of these clones using the T7 and SP6 primers (Table 3.1) was conducted at ILRI (Nairobi, Kenya). The sequences were aligned as described

in Section 3.2.1. Identical sequences were discarded resulting in a final alignment of 24 different sequences.

Table 3.2 PCR conditions used for primer sets detailed in Table 3.1 Components for PCR were added to a master mix as follows: PCR reaction buffer (1x), *Taq* polymerase (0.1 unit), primers (1 µM each), and dNTPs (0.2 mM) and subsequently dispensed into PCR tubes.

	Colony PCR		Genomic PCR	
	T7 and SP6	CP2 specific	Set1	Set 2 and 3
Denaturation	94˚C; 30 s	94˚C; 30 s	94˚C; 1 min	94˚C; 1 min
Annealing	45˚C; 30 s	60˚C; 30 s	55˚C; 1 min	55˚C; 1 min
Elongation	72˚C;1 min	72˚C; 1 min	72˚C; 2 min	72˚C; 2 min
$MgCl_2$ (mM)	2.5	2.5	3	2.5
Reaction volume (µl)	20	20	100	100

3.2.4 Construction of CP PCR library from amplification of genomic DNA

Genomic DNA was isolated from frozen *T. congolense* (strain IL3000) lysate (200 µl) using the Wizard® Genomic DNA Purification Kit. Briefly, lysis solution (400 µl) was added to the sample prior to incubation at 65˚C (5 min). Chloroform (600 µl) was added and mixed. The sample was centrifuged (10 000 *g*, 2 min, RT) and the aqueous phase transferred to a clean tube. Precipitation solution (800 µl) was added and mixed by inversion. DNA was collected by centrifugation (10 000 *g*, 2 min, 4°C) and resuspended in 1.2 M NaCl (100 µl) by vortexing. DNA was precipitated following the addition of 300 µl of 100% (v/v) ethanol at -20˚C for 30 min and dissolved in 50 µl distilled water at 80˚C.

Degenerate primers were designed based upon the alignment of the multiple CP nucleotide sequences of *T. congolense* present in GeneDB (Section 3.2.1). Three sets of primers were designed (Table 2.1 and Figure 2.4): set 1 to amplify the entire 2 kb gene unit based on the MluI restriction site; set 2 to amplify the whole open reading frame (ORF) (i.e. coding for the catalytic and C-terminal domains of CPs); and set 3 to amplify only the sequence coding for the catalytic domain of CPs.

The first primer set was used in the construction of a 2 kb CP PCR library using genomic DNA as a template. The second and third primer sets incorporated cleavage sites for the restriction enzymes NotI and EcoRI in order to facilitate subsequent cloning into an expression vector. Annealing temperatures for PCR were optimised for all three primer sets (Table 3.1 and Table 3.2). A gradient of $MgCl_2$ concentrations (1.5 to 4.0 mM), at a constant annealing temperature (55˚C) was tested for primer set 1. Table 3.2 shows PCR conditions as optimised for all used primer sets.

PCR using primer set 1 for the construction of a 2 kb CP PCR library was conducted using genomic *T. congolense* DNA (at least 50 ng) and performed as summarised in Table 3.2. The PCR product was purified and ligated into the T-vector. An appropriate volume of the ligation mix was used to transform commercial *E. coli* cells by heat shock, and recombinants were selected by blue/white screening and colony PCR as detailed in Table 3.2. Recombinant colonies were used to inoculate 2x YT liquid media containing ampicillin (100 µg/ml) and grown overnight at 37˚C with agitation. The recombinant vectors were purified by miniprep. In order to select clones for sequencing, thirty recombinant clones were screened by PCR using primers specific to congopain (Table 3.1).

The fourteen clones that did not amplify with the congopain specific primers (those as divergent as possible to congopain being of most interest in terms of the present study) were sequenced (ILRI, Nairobi, Kenya) using both the T7 and SP6 primers specific to the cloning vector, and internal primers necessary to obtain the entire 2 kb sequence. The sequences were aligned as described in Section 3.2.1. The construction of the ready-to-express libraries using primer sets 2 and 3 (Figure 3.4), adding NotI and EcoRI cleavage sites, was conducted in a similar manner.

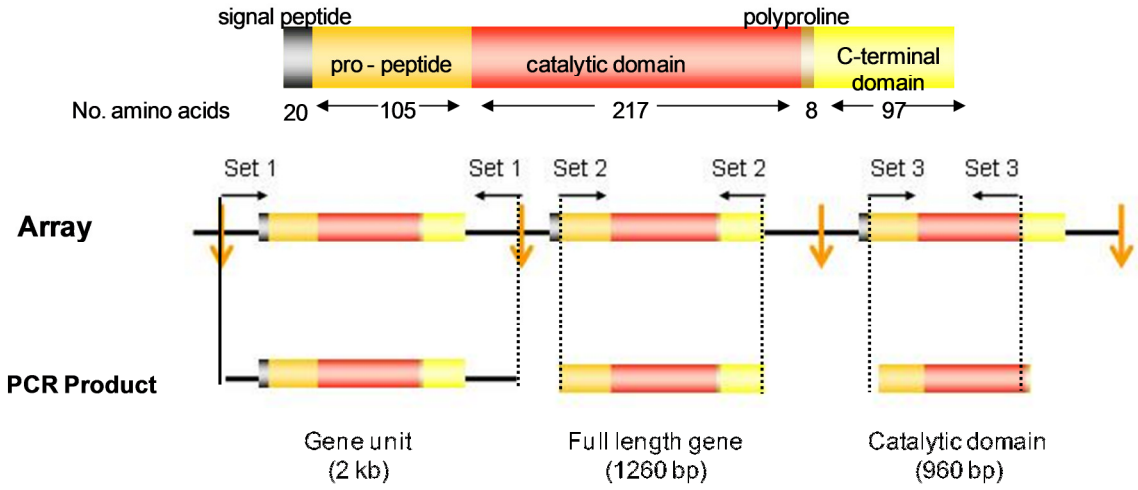


Figure 3.4 Domain structure of congopain and position of primers. Primer sets 1, 2 and 3 were used in the construction of libraries of whole gene units, full length genes and truncated catalytic domain sequences. Primer sets 2 and 3 incorporated sites for the restriction enzymes EcoRI and NotI facilitating subsequent cloning into an expression vector. Vertical yellow arrows indicate the position of the MluI sites in the intergenic regions.

3.3 RESULTS

3.3.1 Two dimensional blotting

The cytoplasmic fraction of *T. congolense* (strain IL3000) proteins was prepared by repeated freeze-thaw cycles and analysed by two dimensional electrophoresis. Two gels were blotted onto nitrocellulose and probed with antibodies specific for (1) congopain-like CPs (mAb 4C5) and (2) all CPs of *T. congolense* (mAb 3B10). Figure 3.5 shows the silver-stained two dimensional separation of the cytoplasmic fraction of *T. congolense*. There are three large high molecular weight spots present as an unresolved smear. The positions of the spots detected by the monoclonal antibodies on the blot are indicated by red (mAb 3B10, detecting all *T. congolense* CPs) and yellow (mAb 4C5, detecting congopain-like CPs of *T. congolense*) circles on Figure 3.5. The actual blots are not shown since the weak signal obtained faded before a photograph could be taken. The four spots recognised by the mAb 3B10, specific for all CPs of *T. congolense* (red circles in Figure 3.5) are in the same horizontal plane. The mAb 4C5, specific for only congopain-like CPs, (yellow circle in Figure 3.5) identified three spots which do not share any common plane. The presence of protein spots on the Coomassie-stained gels after electroblotting confirmed that the blotting process was not totally efficient (data not shown).

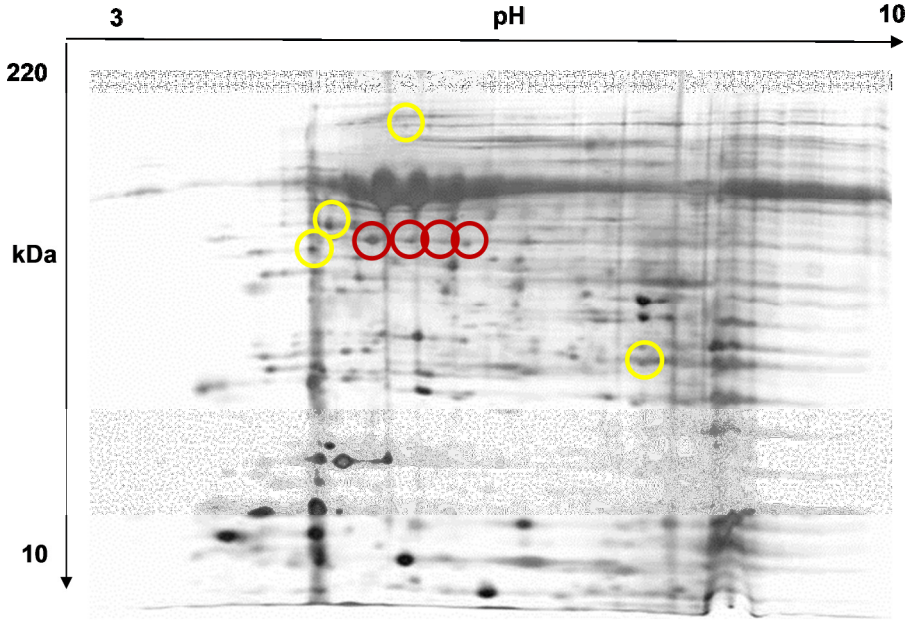


Figure 3.5 Two dimensional electrophoresis of cytoplasmic extract of *T. congolense*. Silver stain of 2D gel. The red circles indicate those proteins detected by western blot with the mAb 3B10 (detecting all CPs of *T. congolense*). The yellow circles indicate those proteins detected by mAb 4C5 (detecting congopain-like CPs of *T. congolense*). The horizontal arrow indicates direction of IEF and the vertical arrow indicates direction of electrophoresis.

3.3.2 CP sub-cosmid library construction

In order to access specific CP genes, two types of libraries were created: (1) a CP sub-cosmid library created by digestion of cosmids containing genomic DNA and (2) a CP PCR library constructed by PCR from genomic DNA. The restriction of a mix of fifteen recombinant cosmids by MluI produced two semi-clear bands at 900 bp and 2 kb and a smear of greater than 5 kb when analysed by agarose gel electrophoresis (Figure 3.6 A). On visual examination of Figure 3.6 A, the smaller band of 900 bp appears to be of greater intensity than the larger 2 kb band. This larger band was excised from the gel and purified for cloning into the pMul vector.

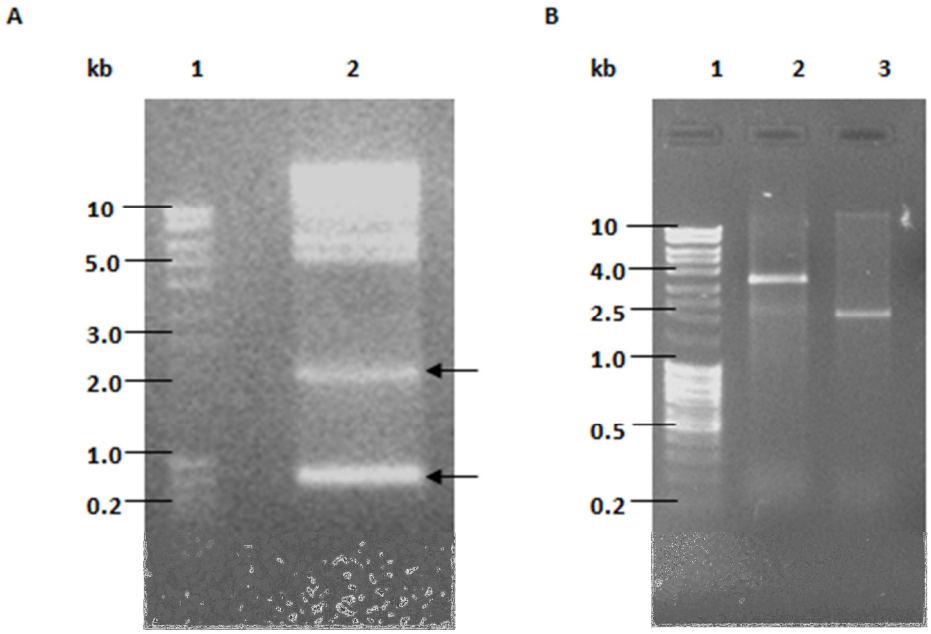


Figure 3.6 Agarose gel analysis of (A) restriction enzyme digest of a mix of fifteen recombinant cosmids containing congopain genes using the restriction enzyme MluI and (B) the pMul vector and the purified 2 kb fragment. (A) Agarose gel [0.8% (w/v)] showing restriction digest. Lane 1: FastRuler® DNA ladder; lane 2: restriction digest (100 µl). Arrows indicate the positions of the 2 kb and 900 bp bands. (B) Lane 1: DNA MassRuler® Mix; lane 2: pMul vector after linearising with MluI; lane 3: 2 kb band excised from agarose gel shown in (A) after purification.

Miniprep DNA of the pMul vector was linearised using the restriction enzyme MluI (Figure 3.6 B, lane 2). The major band in lane 2 corresponds to a size of approximately 3.4 kb, although a faint band at approximately 2.5 kb, likely to correspond to non-linearised vector, is also visible. Lane 3 shows the 2 kb band which was excised and purified from the gel shown in Figure 3.6 A, after MluI digestion. The cohesive ends produced by cleavage of both the pMul vector and purified 2 kb fragment with MluI allowed ligation of the fragment into the vector. Figure 3.6 B shows the linearised pMul vector, which is estimated to be 3.4 kb corresponding to the expected size of the original pGEM®-11Zf+ vector of 3221 bp (Technical Bulletin pGEM-11Zf+ vector, Promega). The band at 2.5 kb may indicate that digestion of the vector was incomplete and the smaller band represents

the "covalently, closed, circular" form, which migrates faster through the gel than the linear DNA due to the tight conformation.

The recombinant vector was subsequently transformed into commercial competent *E. coli* (either K12 or Top10F'). Presence of the *lacZ* gene together with the chromogenic substrate X-gal, allowed blue/white screening for transformant *E. coli* colonies. White (transformant) colonies were selected for colony PCR which was analysed by gel electrophoresis (Figure 3.7). Lanes 2, 4, 6-8, 9, 11, 16 and 18-20 all display a band at 2 kb (absent in other lanes) as well as a smear at the bottom of the gel which is common to all the sample lanes which is likely to be the result of the formation of primer dimers, which are caused by an overabundance of primer in the PCR reaction mix. Lane 14 contains a prominent band at approximately 400 bp which appears to be visible to a lesser extent in most other lanes. Ethidium bromide staining is visible around the wells of lanes 2-20. Presence of the 2 kb band indicated a positive result (i.e. colony was recombinant); therefore, DNA isolated by miniprep from these colonies was sequenced and constitutes the CP sub-cosmid library.

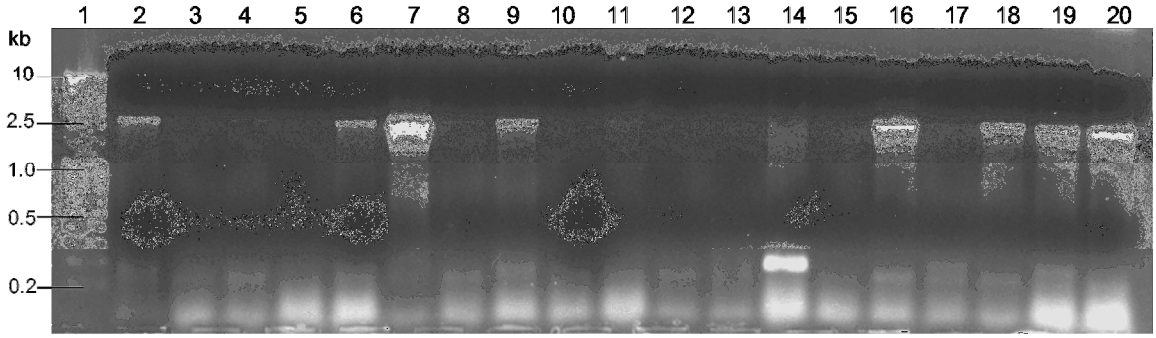


Figure 3.7 Agarose gel analysis of colony PCR for screening for recombinants after transformation with a 2 kb fragment. PCR was performed in a 20 µl volume using T7 and SP6 vector primers. Lane 1: DNA MassRuler® Mix; lanes 2-20: PCR of single white colonies (20 µl).

3.3.3 CP PCR library construction

In a different approach to gain access to individual CP genes, a second library was constructed from PCR of *T. congolense* genomic DNA. Genomic DNA to be used as a template was purified from *T. congolense* (strain IL3000) parasites Figure 3.8 A). A weak band is present higher than 10 kb, out of range of the DNA ladder and a smear is visible at the bottom of the gel, between 100 and 200 bp. The smear at the bottom of the gel represents RNA, which may co-purify with DNA and can be removed by use of an RNAse enzyme such as RNAseH. African trypanosomes have a nuclear genome of approximately 35 Mb which is distributed among three classes of chromosomes: viz. the

diploid megabase chromosomes, intermediate chromosomes of indeterminate ploidy and mini-chromosomes which contain unexpressed VSG genes (Van der Ploeg et al., 1984; El-Sayed et al., 2000).

The annealing temperature used for PCR from genomic DNA using primer sets 1 (whole gene unit, 2 kb), 2 (full-length ORF, 1260 bp) and 3 (catalytic domain, 960 bp) (Table 3.2) was optimised. Figure 3.8 B shows the result of PCR using the three primer sets at annealing temperatures of 45˚C (lanes 2, 4 and 6) and 55˚C (lanes 3, 5 and 7) with 2 mM $MgCl_2$. A smear is present in all sample lanes towards the bottom of the gel, out of range of the lowest band of the DNA ladder (200 bp). Use of primer set 3 at an annealing temperature of 45˚C (lane 2), and primer set 1 at annealing temperatures of both 45˚C and 55˚C (lanes 6 and 7) showed no distinct bands other than the smear. A band of approximately 1200 bp, corresponding to the full-length ORF was amplified by primer set 2 (lane 3). Lanes 4 and 5 each contain a single band of the same size (900 bp, corresponding to the expected size of 960 bp for the catalytic domain), which appears to be more prominent in lane 5 (at an annealing temperature of 55˚C) than lane 4 (45˚C).

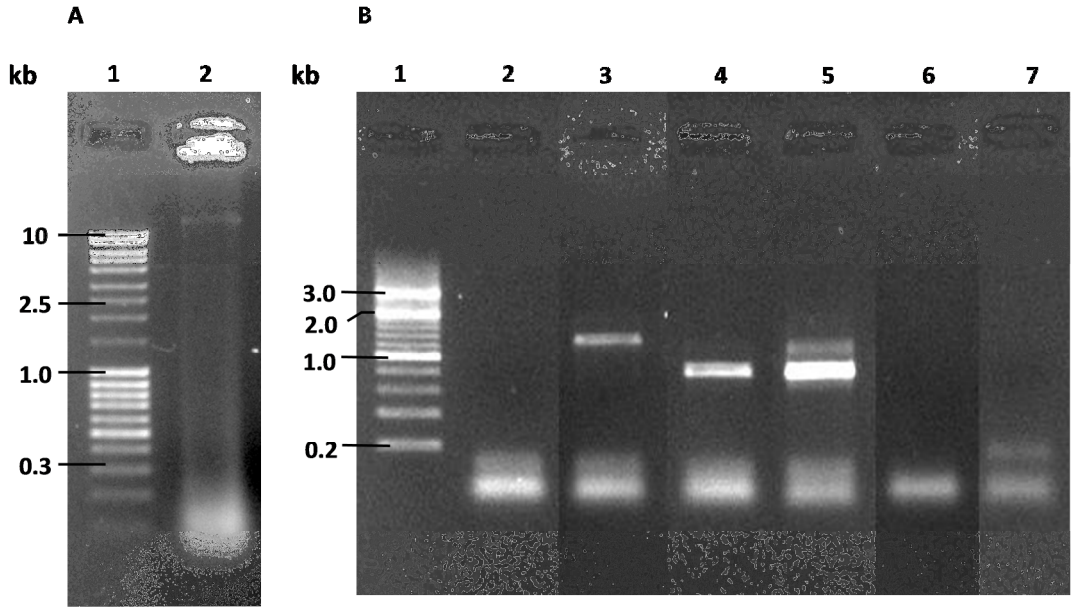


Figure 3.8 Agarose gel analysis of (A) the result of genomic DNA isolation from *T. congolense* (strain IL3000) parasites and (B) optimisation of annealing temperature for primer sets 1, 2 and 3. (A) Lane 1: MassRuler® Mix (1 µl); lane 2: genomic DNA (5 µl). Genomic DNA was purified using the Wizard® Genomic DNA Purification Kit. (B) Lane 1: O'RangeRuler® 200 bp; lanes 2, 4 and 6: PCR using primer sets 3, 2 and 1 respectively at an annealing temperature of 45˚C (10 µl); lanes 3, 5 and 7: PCR using primer sets 3, 2 and 1 respectively at an annealing temperature of 55˚C (10 µl).

To produce a successful PCR for primer set 1 (amplifying the entire 2 kb gene unit), a gradient of $MgCl_2$ concentrations (1.5 - 4 mM) at an annealing temperature of 55˚C was attempted. The result of PCR using these different $MgCl_2$ concentrations is shown in

Figure 3.9. A smear is present in all sample lanes below the lowest band in the DNA ladder (200 bp). Apart from the smear, no other bands are present in lanes 2 and 3, corresponding to $MgCl_2$ concentrations of 1.5 and 2 mM respectively.

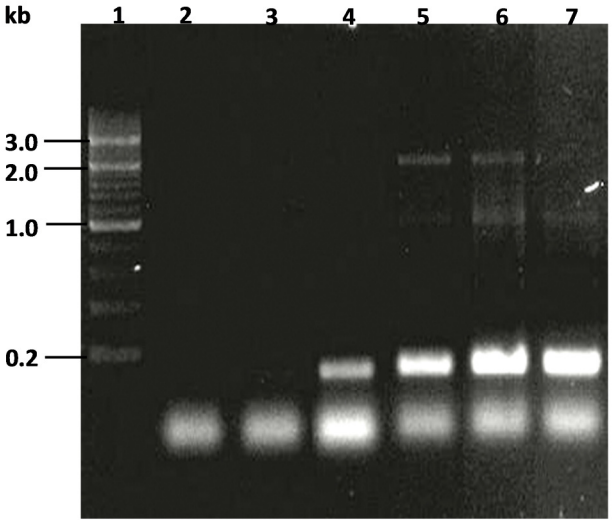


Figure 3.9 Agarose gel showing optimisation of $MgCl_2$ concentration for primer set 1 amplifying the whole 2 kb gene unit. Lane 1: O'RangeRuler® 200 bp; lanes 2-7: PCR with set 1 primers at $MgCl_2$ concentrations of 1.5, 2.0, 2.5, 3.0, 3.5 and 4.0 mM respectively. Genomic DNA was purified using the Wizard® Genomic DNA Purification Kit.

Lanes 4 - 7 all contain a band at approximately 200 bp, which increases in intensity as the $MgCl_2$ concentration increases. In lane 5 (3.0 mM $MgCl_2$), a faint band of 2 kb is present. This 2 kb band is also present in lanes 6 and 7 (3.5 and 4.0 mM $MgCl_2$ respectively), together with a smaller band at approximately 1200 bp. These two bands are more distinct in lane 6 than lane 7. PCR products for the three primer sets were purified from the gel, as shown in Figure 3.10.

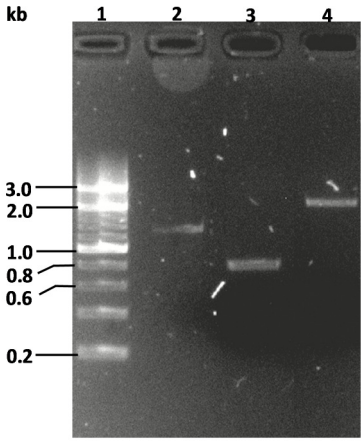


Figure 3.10 Agarose gel showing purified PCR products from primer sets 1, 2 and 3 used for cloning into a T-vector. PCR products were purified from the gel using the Wizard® PCR Preps Kit. Lane 1: O'RangeRuler® 200 bp; lane 2: primer set 3 amplicon; lane 3: primer set 2 amplicon; lane 4: primer set 1 amplicon.

Lane 2 (primer set 2 amplifying the full length ORF) shows a single band at 1200 bp. A band at approximately 900 bp is present in lane 3 (primer set 2 amplifying the catalytic domain), while a single band at 2 kb is evident in lane 4 (primer set 1, amplifying the entire gene unit). The purified PCR products for all three primer sets, once ligated into a T-vector, were inserted into *E. coli* by transformation. As before, blue/white screening for transformants was possible, allowing transformant colonies (white) to be selected for recombinant screening by PCR using the T7/SP6 T-vector primers amplifying the multiple cloning site of the T-vector.

3.3.4 Analysis of CP sub-cosmid and PCR library sequences

The information obtained from sequencing random clones from the 2 kb CP sub-cosmid and CP PCR libraries is summarised in Table 3.3. An overview of the variation at significant areas of the peptidase sequences is given, i.e. the "YHNGA" pro-peptide motif which, according to Lalmanach *et al.* (1998), specifically inhibits activity of congopain; the N-terminal region of the catalytic domain; the catalytic triad residues and the polyproline hinge region between the catalytic and C-terminal domains.

Table 3.3 A summary of the variation of congopain-like genes obtained from CP sub-cosmid and PCR libraries compared to database reads. The 2 kb CP sub-cosmid library was created by cloning a 2 kb fragment obtained from a cosmid library containing digested *T. congolense* genomic DNA. The PCR library was created using PCR with degenerate primers to amplify directly from *T. congolense* genomic DNA. Sequences were obtained from the *T. congolense* read database hosted at GeneDB (http://www.genedb.org/, accessed 7-5-2008) by a BLASTN search using the congopain nucleotide sequence as a query.

	2 kb CP sub-cosmid library	2 kb CP PCR library	Database reads (BLASTN with CP2)
Number of sequences analysed	24	22	29
"YHNGA" motif	17/17	18/21	23/24*
N-terminal residues			
APEA	12/24	17/21	17/25*
APPA	0	0	6/25*
APDA	2/24	0	0
PPMT	10/24	4/21	2/25*
Catalytic triad (papain: $C^{25}H^{159}N^{175}$)			
CHN	1/24	11/21	22/29
SHN	5/24	0/21	1/29
SSN	14/24	7/21	4/29
SFN	0/24	1/21	0/29
SYN	0/24	1/21	0/29
FHN	3/24	1/21	2/29
FYN	1/24	0	0

*Reads obtained from the databases vary in length; therefore, not all of the sequences include the entire gene.

It should be noted that the "APDA" N-terminal motif is not found at all in the database sequences but is present in the 2 kb sub-cosmid library; conversely, the "APPA" N-

terminal motif is present in the database sequences but not in the 2 kb sub-cosmid library. Concerning the catalytic triad, the majority of the genes in the read database have the peptidase catalytic triad of Cys^{25}, His^{159} and Asn^{175} (papain numbering), whereas the majority of the sub-cosmid library genes have a Ser^{25}, Ser^{159} and Asn^{175} catalytic triad.

3.4 DISCUSSION

Two dimensional electrophoresis of the *T. congolense* cytoplasmic proteome was conducted in order to visualise expression of isoforms of congopain. The large spots of high molecular weight present on the gel (Figure 3.5) correspond to the variable surface glycoprotein (VSG) of the parasite at approximately 53 kDa (Lalor *et al*., 1984). The VSG protein is responsible for the phenomenon of antigenic variation in bloodstream forms of the parasite (Barry and Carrington, 2004). Since only cytoplasmic proteins were intended for extraction using the process outlined before (Section 2.2.2), the VSG protein can be considered a contaminant since the VSG should be part of the membrane fraction rather than the cytoplasmic fraction. It has, however, been observed that *T. congolense,* like *T. equiperdum* and *T. evansi,* contains high levels of phospholipase C (Lamont *et al*., 1987). Upon osmotic lysis, this endogenous phospholipase C is able to cleave membrane-bound VSG from the GPI anchor. This releases VSG from the GPI anchor and consequently from the membrane, giving rise to soluble VSG. This could explain the presence of VSG in the cytoplasmic extract.

The mAb 3B10 has been previously raised against the recombinantly expressed catalytic domain of congopain, C2 (Boulangé et al., 2011), and thus far, the specific epitope that the mAb interacts with is unknown, although it is thought to be a conformational epitope since the antibody does not recognise the reduced form of the protein (E. Authié and T. Lefrançois, unpublished observation). Monoclonal antibody 4C5 was raised against the entire native congopain purified by electroelution (Authié *et al*., 1992), but, given the high immunogenicity of the C-terminal region, and the fact that it does not recognise the recombinant catalytic domain, it is likely that mAb 4C5 recognises an epitope in the C-terminal domain (Boulangé *et al*., 2001). In addition, mAb 4C5 has been found not to react with CP1 (A. Boulangé, unpublished observation) which is expected given the sequence differences between the CP1 and congopain in the C-terminal region. The proteins detected by mAb 3B10 have the same molecular weight, since all are at the same horizontal plane, but different pI-values as shown by their distribution in the vertical plane. For unknown reasons, the molecular weight markers are not visible on the gel; it is, therefore, difficult to ascertain the size of the proteins recognised by the antibodies.

Still, the position of the VSG (53 kDa) allows an estimation of the size of the proteins recognised by mAb 3B10. The detection of proteins of around 30 kDa by mAb 3B10 indicates that these proteins are probably isoforms of congopain or other, related, CPs. The identification of the group of proteins by mAb 3B10 implies that variable forms of congopain are present at the protein level.

Although it has been established that at least one other expressed congopain-like gene does exist, native CP1 protein has never been isolated from bloodstream forms of the parasite (Fish *et al.*, 1995). In addition, work has been done (A. Boulangé, unpublished) showing the presence of mRNAs encoding several different CP genes. However, since trypanosome genes are generally polycistronic, i.e., several genes are under the control of a single promoter, transcription of the whole congopain gene array is likely to occur at the same time generating a single mRNA transcript encoding the whole congopain gene array (Pays, 2005; Gomez et al., 2010). Due to this polycistronic nature of the trypanosome genome, it is likely that regulation of gene expression occurs at the mRNA level rather than transcriptional control (Hotz *et al.*, 1997) and there is evidence for this in *T. cruzi* (Tomás and Kelly, 1996). For this reason, it is impossible to infer protein expression solely by analysis of mRNA and, consequently, visualisation of the protein itself is the only way to verify expression. Therefore, this tentative identification of multiple forms of congopain-like proteins, as identified by mAb 3B10 on a blot of the two dimensional separation of *T. congolense* proteins, is viewed to be significant.

Possible future work concerning two dimensional electrophoresis includes the optimisation of the transfer process for large gels, using different buffer systems, different blotting parameters and comparing semi-dry blotting to wet blotting. Also, the removal of the VSG from the cytoplasmic fraction would improve the visualisation since the VSG is large and abundant and may be obscuring other proteins. In addition, the large quantity of VSG means that the protein sample consists mostly of VSG. Therefore, if the VSG can be removed, more proteins of interest can be loaded onto the gel, allowing better visualisation and blotting. Possible strategies to remove the VSG include affinity chromatography or immunoprecipitation using anti-VSG antibodies. Also, phospholipase C inhibitors could be used in order to prevent the production of soluble VSG (Stanton *et al.*, 2002), thus keeping VSG out of the soluble protein fraction.

Alternatively, the CPs could be purified from the total cytoplasmic protein extract using the mAbs 3B10 and 4C5, either by affinity chromatography or immunoprecipitation. This

method might be preferable since more proteins of interest could be loaded onto the gel, and a smaller gel could be used facilitating downstream processes such as blotting. Once the two dimensional electrophoresis process would have been optimised, further study of the proteins of interest by mass spectrometry (MALDI-ToF analysis) or N-terminal sequencing could be attempted.

As described in the introduction, a MluI cleavage site is present in the highly conserved intergenic region between the genes in the congopain-like gene array (Figure 3.2). Since each gene unit is approximately 2 kb in size, cleavage of the tandem gene array by MluI should release individual gene units of 2 kb each and this is represented by the band at 2 kb in Figure 3.6 A. In addition to the MluI restriction enzyme site in the intergenic region, the congopain gene has a MluI site in the catalytic domain; therefore, the MluI digestion of congopain within the gene array would yield a fragment smaller than 2 kb (approximately 900 bp as calculated from Figure 3.11).

The band smaller than the 2 kb band on the gel in Figure 3.6 A corresponds to a length of approximately 900 bp, therefore, it can be concluded that this band is the result of digestion of both MluI sites in the congopain gene unit. Analysis of sequences from the *T. congolense* database (see Table 3.3) indicated that those genes which translate with a catalytic triad of CHN, like congopain itself, are the only ones in the array to contain the additional MluI site in the catalytic domain. For this reason, digestion of the recombinant cosmids using MluI and cloning of only the 2 kb fragment results in selection of those genes with least sequence homology to congopain, and therefore, is of most interest in terms of the present study.

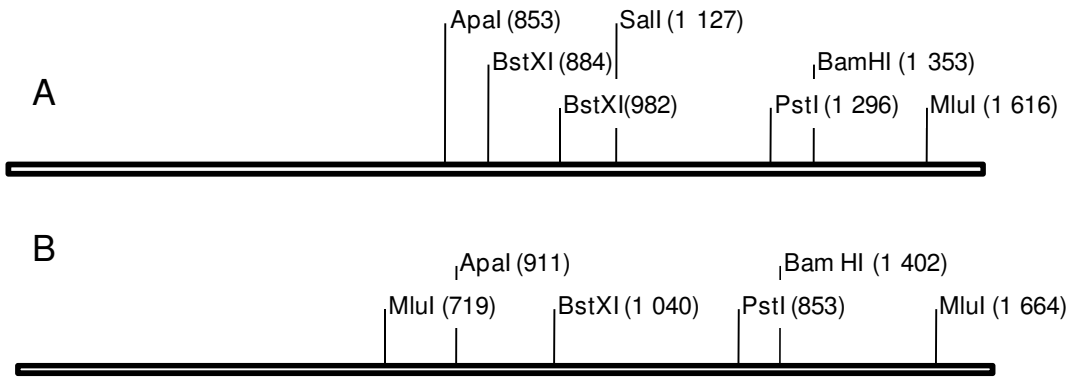


Figure 3.11 Restriction enzyme map of DNA from (A) CP1 and (B) CP2 (congopain). Map generated using Sequencher® software (Gene Codes Corporation, 2006). Sequences obtained from GenBank® (http://www.ncbi.nlm.nih.gov/genbank/, accessed 5-5-2008), GenBank ID Z25813 and L25130 respectively.

Sequence differences of the CP genes from the different sources (sub-cosmid and PCR libraries and databases) were analysed and some significant regions of the sequences are highlighted in Table 3.3. One of the observations which can be made from these sequences comparisons is that the "YHNGA" peptide, found in the pro-peptide region of congopain, is highly conserved. The significance of the conservation of this motif stems from the fact that, like the mammalian cathepsins, congopain is secreted as an inactive zymogen (procongopain) and is activated by autocatalytic cleavage of the pro-peptide region (Serveau *et al.*, 2003).

It has been established that the pro-peptides of mammalian cathepsins K, L and S efficiently inhibit their respective enzymes usually acting as slow-binding competitive inhibitors and, furthermore, are selective for cathepsin L-like enzymes showing no significant inhibition of cathepsin B (Carmona et al., 1996; Billington et al., 2000; Guay et al., 2000; Mendoza-Palomares et al., 2008). Conversely, the propeptide of cathepsin B, while a potent inhibitor of the parent enzyme, did not inhibit papain. The same phenomenon of pro-peptide inhibition of cathepsin-like enzymes has been observed with peptidases from the parasites *F. hepatica*, *T. cruzi* and *T. congolense* (Lalmanach et al., 1998; Roche et al., 1999; Reis et al., 2007). Of particular significance is the fact that the pro-peptide of the cathepsin L-like enzyme from *F. hepatica*, was selective for only the parent enzyme and did not appreciably inhibit mammalian cathepsins L, K or S indicating that the pro-peptide could be a potentially useful selective inhibitor of the parasite enzyme (Roche *et al.*, 1999).

For *T. congolense*, Lalmanach *et al.* (1998) showed that a 15-residue peptide, based on the conserved "YHNGA" motif in procongopain, inhibited the proteolytic activity of congopain and cruzain (which also contains the motif), but not mammalian cathepsins B and L. Due to the presence of the "YHNGA" peptide in almost all sequences examined, an inhibitor based upon this motif is likely to inhibit all congopain-like CPs. More recent findings, however, show that the recombinantly expressed full length pro-domain of cruzain inhibited not only the trypanosomal CPs cruzain and brucipain (*T. b. brucei*), but also human cathepsins F and V (Reis *et al.*, 2007).

The CP1 gene, described by Fish *et al.* (1995) was cloned and subsequently sequenced from a cDNA library of *T. congolense*, and was characterised by an APPA N-terminal cleavage site, a CHN catalytic triad and a hinge region with eight proline residues. CP1 is not present in either the sub-cosmid or PCR libraries prepared in this study but is found

within the database sequences (as retrieved by BLASTN search of GeneDB *T. congolense* read sequences from http://www.genedb.com, accessed 5-5-2008). The methods employed in the construction of the sub-cosmid and PCR libraries were both based on the presence of the MluI site in the intergenic region. Therefore, the absence of CP1 from the libraries prepared in this study may be due to the fact that the CP1 gene may not be flanked by MluI cleavage sites i.e. it is not located within the congopain multi-gene array but elsewhere in the genome.

Another of the significant findings from the sequencing of the sub-cosmid and PCR libraries is the fact that the majority of the database sequences have the expected CP triad, i.e. Cys^{25}, His^{159} and Asn^{175} (papain numbering), whereas the majority of the 2 kb CP sub-cosmid library genes have a Ser^{25}, Ser^{159} and Asn^{175} catalytic triad. In fact, only a single gene from the 2 kb CP sub-cosmid library possessed the conventional CHN triad. These results indicate that the cloning of the 2 kb fragment from the MluI digested recombinant cosmids did indeed select for those genes with the least sequence homology to congopain. The sequences obtained by PCR from genomic DNA (second column Table 3.3) are similar to what is obtained from the database (third column), implying that this is likely to be the actual proportion of the different types of genes present, as opposed to the sequences obtained from the sub-cosmid library (first column, Table 3.3). The sequences obtained from the two libraries prepared in this study, together with the analysis of the congopain-like genes present in the databases, clearly indicate that there is significant variability at the nucleic acid level within the congopain-like genes present in *T. congolense.* It is necessary to further investigate the congopain variants by cloning and recombinant expression to compare the enzymatic characteristics of the variant enzymes.

CHAPTER 4: RECOMBINANT EXPRESSION AND CHARACTERISATION OF VARIANT CONGOPAIN-LIKE CYSTEINE PEPTIDASES OF *TRYPANOSOMA CONGOLENSE*

D. Pillay[1*], A. Boulangé[2] and T.H.T. Coezter[1]

[1] School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal (Pietermaritzburg campus), Private Bag X01, Scottsville 3209, South Africa
[2]Trypanosomoses, UMR-IRD/CIRAD 017, CIRAD-Emvt, 34398 Montpellier, France
[*] Email: 202511575@ukzn.ac.za

Congopain, the major cysteine peptidase of *Trypanosoma congolense* is an attractive candidate for an anti-disease vaccine and target for the design of specific inhibitors. A complicating factor for the inclusion of congopain in a vaccine is that multiple variants of congopain are present in the genome of the parasite. In order to determine whether the variant congopain-like genes code for peptidases with enzymatic activities different to those of congopain, two variants were cloned and expressed. Two truncated catalytic domain variants were recombinantly expressed in *Pichia pastoris*. The two expressed catalytic domain variants differed slightly from one another in substrate preferences and also from that of C2 (the recombinant truncated form of congopain). Surprisingly, a variant with the catalytic triad Ser^{25}, His^{159} and Asn^{175} (papain numbering) was shown to be active against classical cysteine peptidase substrates and inhibited by E-64, a class specific cysteine peptidase inhibitor. Both catalytic domain clones and C2 had pH optima of either 6.0 or 6.5 implying that these congopain-like peptidases are likely to be expressed and active in the bloodstream of the host animal.

A cysteine peptidase (CP), named congopain, has been shown to act as a pathogenic factor in *Trypanosoma congolense*-infected cattle (Authié *et al*., 2001). In addition, antibodies against congopain correlate with resistance to the disease (trypanotolerance) (Authié *et al*., 1993a). For these reasons, congopain is a definite target for an anti-disease vaccine or synthetic inhibitors.

Congopain, however, like the cathepsin L-like peptidases of *Leishmania mexicana* and *T. cruzi,* appears to be present in multiple isoforms (Mottram et al., 1997; Lima et al., 2001; Caffrey and Steverding, 2009). Indeed, CP1 (GenBank® ID: Z25813), a divergent congopain-like peptidase has been described in *T. congolense* (Fish et al., 1995; Boulangé et al., 2001), but unlike congopain, native CP1 has not been isolated from the parasite to date. For the development of a vaccine or inhibitors targeting congopain, a study of the variant congopain-like enzymes would be required.

In a study aimed at elucidating the genomic organisation of congopain genes in a given *T. congolense* clone (strain IL3000) (Wellde *et al*., 1974), Kakundi (2008) showed by Southern blotting experiments that congopain-like CPs exist as a multigene family of 2 kb

gene units organised as a tandem array (see Figure 3.2). This study was followed by the construction of a cosmid library of 105 clones of 40 - 50 kb fragments of *T. congolense* genomic DNA. Analysis of these recombinant cosmids indicated that each gene unit is 2 kb in length, consisting of a coding region and a highly conserved intergenic region and the sequence homology between the genes ranges from 86 – 99% (Kakundi, 2008).

Interestingly, many of the genes display substitutions of the active site residues. The classical CP catalytic triad is Cys^{25}-His^{159}-Asn^{175} corresponding to Cys^{25}- His^{191}-Asn^{211} in congopain (Lalmanach et al., 2002; Lecaille et al., 2002; Barrett and Rawlings, 2004b). For this reason it was surprising to note that one of the most often recurring substitutions is that of the active site cysteine (Cys^{25}) for a serine (Ser^{25}). Also, the active site His^{191} (congopain numbering, Figure 4.1) was less often substituted by a Ser or Tyr. According to the methods used to estimate the prevalence of the genes bearing these substitutions, it was found that 40-60% of the congopain-like genes present this type of mutation (see Table 3.3). This phenomenon is puzzling, and therefore, it was of interest to determine whether these variant genes were active to perform a function distinct to that of the characterised cysteine peptidase, congopain.

In the present study, the catalytic domains of two variant CPs, named CP_{SYN} and CP_{SHN} (named according to the respective catalytic triads) were expressed, purified and characterized enzymatically in comparison to C2, the recombinant catalytic domain of congopain.

4.1 MATERIALS

General molecular biology: SacI, SalI, PmeI, EcoRI, NotI, shrimp alkaline phosphatase (SAP), T4 DNA ligase, dNTP mix, X-gal, MassRuler® DNA Ladder Mix, GeneRuler® DNA ladder Mix, FastRuler® DNA ladder, O'RangeRuler® DNA Ladder 200 bp and the GeneJet® Plasmid Miniprep Kit were obtained from Fermentas (Vilnius, Lithuania). *Taq* polymerase was obtained from Solis Biodyne (Tartu, Estonia). The Wizard® PCR Preps DNA Purification System, the Wizard® Plus SV Minipreps DNA Purification System, the Wizard® Genomic DNA Purification Kit, and the pGEM-T® Easy cloning kit were purchased from Promega (Madison, WI, USA). The DNA Clean and Concentrator Kit® was purchased from Zymo Research (Orange, CA, USA).

***E. coli* cells:** Competent *E. coli* cells (NEB 5-alpha K12 and One Shot® Top10F') were purchased from New England Biolabs (Ipswich, MA, USA) and Invitrogen (Carlsbad, CA, USA) respectively. Both *E. coli* strains used allow blue/white screening for transformants in the presence of X-gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) by alpha-complementation of the beta-galactosidase gene.

***P. pastoris* expression:** The *P. pastoris* expression system including the yeast strains X-33 and GS115, the vector pPic9 and the antibiotic Zeocin® were purchased from Invitrogen. The Bio-Rad Gene Pulser® used for electroporation was obtained from Bio-Rad (Hercules, CA, USA). The modified pPicZαA28 vector was a gift from Prof. T. Baltz (University Victor Segalen, Bordeaux, France).

Protein purification: Filter paper was obtained from Whatman (Middlesex, UK). DEAE-650M anion exchange resin was purchased from Tosoh Bioscience (Tokyo, Japan). Sephacryl S300 HR resin was obtained from Sigma (Munich, Germany). Centricon® centrifugal concentrators were purchased from Millipore (Billerica, MA, USA).

Antibodies: IgY anti-congopain-N-terminus peptide antibody (affinity purified, 50 µg/ml) was raised against the 22 N-terminal residues of congopain (Mkhize, 2003). Monoclonal antibodies 4C5 (Authié *et al*., 1992) and 3B10 (T. Lefrançois, unpublished, CIRAD, Montpellier, France), raised against native congopain and the recombinantly expressed catalytic domain of congopain (C2) respectively, were obtained from A. Boulangé (Boulangé et al., 2011). Goat anti-mouse IgG HRPO conjugate and rabbit anti-IgY HRPO conjugate were obtained from Sigma (Munich, Germany).

Peptide substrates and inhibitors: The peptide substrates H-D-Ala-Leu-Lys-aminomethyl coumarin (AMC), H-Ala-Phe-Lys-AMC and H-Pro-Phe-Arg-AMC were obtained from Bachem (King of Prussia, PA, USA). Benzyloxycarbonyl (Z)-Phe-Arg-AMC, *trans*-epoxysuccinyl-L-leucyl-amido(4-guanidino)butane (E-64), leupeptin, pepstatin and 4-(2 –Aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF) were purchased from Sigma. Fluorescence assays were measured using a FLUOR Star Optima spectrophotometer from BMG Labtech (Offenburg, Germany).

4.2 METHODS

4.2.1 Cloning and expression of CP catalytic domains in *P. pastoris*

The two CP catalytic domain sequences were amplified using a forward primer *T. congolense* (strain IL3000) (5' GCC GAA TTC GCG GC TYT GTT CCC GTG GCG 3') and a reverse primer for the truncated catalytic domain (5' TAG CGG CCG CAC CCT CGC ASA ACT CCT GCG TGA 3') resulting in a 960 bp product. Underlined letters indicate restriction sites for EcoRI and NotI inserted into the primer for facilitating sub-cloning into the expression vector. Amplicons were cloned into pGEMT® vector and transformed into competent *E. coli* JM109 by heat-shock. The restricted EcoRI-NotI fragments were sub-cloned into the *P. pastoris* expression vector pPic9. Following linearization of the recombinant vectors with SacI or PmeI, the recombinant vectors were transformed into non-recombinant *P. pastoris* strain GS115 and grown on selective media (Wu and Letchworth, 2004).

Yeast extract peptone dextrose (YPD) [1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose] liquid media (usually 50 ml) was inoculated with a single colony and grown at 30˚C in baffled flasks with agitation for two days. The pre-culture was used to inoculate a larger volume (usually 500 ml) of Buffered Media Glycerol Yeast (BMGY) [1% (w/v) yeast extract, 2% (w/v) peptone, 100 mM potassium phosphate buffer pH 6.5, 1.34% (w/v) yeast nitrogen base without amino acids) that was grown for a further three days. The cells were collected by centrifugation (2000 *g*, 10 min, 4˚C) and resuspended in an equal volume of buffered minimal media [BMM, 100 mM potassium phosphate buffer, pH 6.5, 1.34% (w/v) yeast nitrogen base without amino acids, 0.0004% (w/v) biotin, 5% (v/v) methanol]. Once in BMM, culture flasks were covered with three layers of sterile cheesecloth to facilitate aeration and methanol [0.5% (v/v)] added daily for the duration of expression (usually four days).

A C2-pPic 9 recombinant plasmid (Boulangé et al., 2011), containing the gene coding for the catalytic domain of congopain, was transformed and C2 expressed alongside the two CP variants as a positive control. The expression protocol described above was developed after monitoring expression while changing several variables including the use of Buffered Media Methanol Yeast (BMMY) [BMGY with 1% methanol instead of glycerol] as an expression medium, buffering pH of the BMM (pH 6.0 – 7.5), expression time (2 – 7 days) and amount of methanol for induction (0.05 – 0.5%). Supernatants were analysed by SDS–PAGE (Laemmli, 1970) and western blotting using affinity purified chicken anti-

congopain N-terminal peptide IgY (1 µg/ml) incubated for 2 h at RT. The nitrocellulose membrane was incubated with rabbit anti-IgY HRPO conjugate (1:12 000) (Mkhize, 2003) followed by substrate solution [0.06% (w/v) 4-chloro-1-naphthol, 0.1% (v/v) H_2O_2].

4.2.2 Purification of catalytic domain variants

Cultures were centrifuged (2000 *g*, 10 min, 4˚C) and supernatant fluids collected. If not immediately used, the supernatant fluids were stored in plastic vessels at -20˚C. Three phase partitioning (TPP) described by Pike and Dennison (1989) was used as an initial concentration method. Briefly, the supernatant fluids (usually 500 ml) were filtered (Whatman No. 4 filter paper) and tert-butanol [30% (v/v) final] added. Ammonium sulfate [30% (w/v)] was added and stirred until completely dissolved. This mixture was centrifuged (6000 *g*, 10 min, 4˚C) in a spin-out rotor so that the three phases were completely delineated. The tert-butanol and aqueous phases were removed leaving the protein precipitate. The protein precipitate was dissolved in a minimal volume of sterile water and dialysed against at least two changes of phosphate buffer (50 mM, pH 7.2) at 4˚C. The dialysed sample was concentrated against polyethylene glycol (M_r 20 000) until the final volume was approximately 5 - 10 ml.

The protein samples were further purified by DEAE - 650M anion exchange chromatography. The column (2.5 x 10 cm) was equilibrated with 10 mM Tris–Cl buffer pH 8.0 at a flow rate of 30 $cm.hr^1$. Following loading of the protein sample the column was washed with Tris–Cl buffer until the absorbance at 280 reached baseline and bound protein eluted with a linear gradient of NaCl from 0 to 1 M (5 col. vols). Elution of protein was monitored by reading the absorbance of each 1.5 ml fraction collected at 280 nm. Enzyme-containing fractions identified by measuring the hydrolysis of Z-Phe-Arg-AMC were pooled and concentrated using Centricon centrifugal concentrators according to the manufacturer's instructions.

Molecular exclusion chromatography (MEC) was conducted using a Sephacryl S300 HR column (25 x 840 mm, 25 $ml.h^{-1}$, 4˚C), calibrated using 2 mg/ml blue dextran and 5 mg/ml each of BSA, ovalbumim, carbonic anhydrase, soybean trypsin inhibitor and lysozyme. The column was equilibrated with MES buffer (0.5 M, pH 6.5) and sample (approximately 5 ml) applied. Peptidase-containing fractions identified by measuring the absorbance at 280 nm were pooled and concentrated using Centricon® centrifugal concentrators according to the manufacturer's instructions. Protein concentration was determined using the bicinchoninic acid (BCA) protein assay kit (Pierce). Purified samples were analysed

by reducing SDS-PAGE (Laemmli, 1970) and protein bands visualised by silver staining (Blum *et al.*, 1987). Recombinant C2 was purified in the same manner and used as a control in the enzymatic analysis and inhibition experiments.

4.2.3 Enzymatic analysis of catalytic domain variants

Proteolytic activity of C2 and the CP variants was analysed on gelatin containing non-reducing SDS-PAGE gels (Heussen and Dowdle, 1980), followed by incubation in congopain assay buffer [100 mM Bis-Tris, pH 6.0, 4 mM EDTA, 0.02% (w/v) NaN_3, 6 mM DTT]. The concentration of active protein was determined by active site titration using the irreversible cysteine peptidase inhibitor E-64 (Barrett *et al.*, 1982) and the substrate Z-Phe-Arg-AMC. Stock solutions of E-64 (1 mM in water) and Z-Phe-Arg-AMC [50 mM in dimethyl sulfoxide (DMSO)] were prepared. Briefly, enzyme (0.1 - 1 µM) diluted in 0.1% (v/v) Brij-35 was incubated with E-64 (0.1 - 1 µM) in congopain assay buffer [100 mM Bis-Tris buffer, pH 6.0, 4 mM EDTA, 0.02% (w/v) NaN_3, 6 mM DTT] at 37˚C for 30 min; and activity against Z-Phe-Arg-AMC (20 µM) measured (excitation 360 nm; absorbance 460 nm) in a FLUOR Star Optima spectrophotometer. A standard curve for the AMC product was prepared by diluting AMC (50-250 pmol) in assay buffer. The linear regression of the calibration curve allowed quantitation of fluorescence allowing measurement of activity in pmol of AMC per unit time (s). All enzyme assays and the calibration curve were conducted in standard 100 µl volumes.

Activity assays on active site-titrated enzyme were conducted using Z-Phe-Arg-AMC substrate. Briefly, enzyme (1.5 ng active enzyme) diluted in 0.1% Brij was incubated with congopain assay buffer containing 6 mM DTT for 10 min at 37˚C. Substrate (0 - 100 µM) was added and fluorescence measured (excitation 360 nm; absorbance 460 nm) for thirty cycles. Michaelis-Menten plots were used to determine the kinetic constants K_m and V_{max} using the software Hyper32® (1991-1992, J.S. Easterby, University of Liverpool, UK).). The specificity constant, k_{cat} was calculated from the formula $k_{cat} = V_{max}/E_o$, where E_o is the concentration of active enzyme (Salvesen and Nagase, 2001). Similar assays were conducted for the substrates Z-Pro-Arg-AMC, H-D-Ala-Leu-Lys-AMC and H-Ala-Phe-Lys-AMC which were all prepared as 50 mM stocks in DMSO.

In order to determine the optimum pH for each congopain CP variant, a pH profile was generated. Constant ionic strength AMT buffers (100 mM Na-acetate, 200 mM Tris-Cl, 100 mM MES, 6 mM DTT, 4 mM Na_2EDTA) were prepared from pH 4.0 to pH 9.0 in 1.0 increments by titration with NaOH or HCl (Ellis and Morrison, 1982). Enzymes (1.5 ng

active enzyme) diluted in 0.1% (v/v) Brij-35 were incubated in each buffer of a different pH for 5 min at 37˚C. The substrate H-Ala-Leu-Lys-AMC (20 µM, 25 µl) was added and fluorescence measured (excitation 360 nm; absorbance 460 nm).

Inhibition assays were performed on the recombinant enzymes. For E-64 (1 mM, water) leupeptin (1 mM, water), AEBSF (100 mM, water) and pepstatin A (1 mM, methanol), the percentage of inhibition was estimated by measuring residual enzyme activity. Briefly, the enzyme (2 ng, 50 µl) diluted in 0.1% (v/v) Brij-35 was activated in congopain assay buffer containing 6 mM DTT (25 µl) at 37˚C for 5 min. The inhibitor was added (50 µM, 25 µl) and incubated with the enzyme at 37˚C for 15 min. Residual activity was measured using H-Ala-Leu-Lys-AMC (20 µM). For the diazomethyl-ketone (DMK) inhibitors (prepared as 100 mM stocks in DMSO), the second order rate constant (k_{ass}) was measured. Enzyme (10 nM) was activated in congopain assay buffer containing 6 mM DTT at 37˚C for 5 min. Inhibitor (10 nM) was added and the mix subsequently assayed against H-Ala-Leu-Lys-AMC (20 µM). The observed association rate constant (k_{obs}) was obtained from the slope of a plot of the natural log of residual activity against time. The true association constant, k_{ass} and time required for free enzyme concentration to decrease by 50% ($t_{½}$) was calculated according to Salvesen and Nagase (2001). The apparent inhibition constant $K_{i(app)}$ was calculated from the following equation: $V_o / V_i = 1 + [I] / (K_{i(app)})$; where $[I]$ denotes the inhibitor concentration. True K_i could subsequently be derived from: $K_i = (K_{i(app)}) / 1 + ([S]/K_m)$.

4.3 RESULTS

4.3.1 Sequence alignment of C2 and Cp variants

An alignment of the translated catalytic domains of the expressed variants is shown in Figure 4.1. The “YHNGA” pro-peptide motif (residues 5-9 in Figure 4.1) which has been shown to specifically inhibit the activity of congopain (Lalmanach et al., 1998; Godat et al., 2005) is conserved in all three sequences. The processing of congopain from inactive proenzyme to the active form is achieved by cleavage of the pro-peptide at the “TGK↓APEA” (residues 27-33 in Figure 4.1) (Serveau *et al.*, 2003).

For this reason it is surprising that the N-terminal residues of the mature CP_{SHN} and CP_{SYN} enzymes are different, “APDA” and “PPDA” respectively. It is important to note that the change of the catalytic cysteine to a serine (residue 54 in Figure 4.1) is the result of two base pair changes i.e. the codon changes from “TGC” (cysteine) in C2 to “TCG” (serine)

in CP_{SHN} and CP_{SYN}. Similarly, the change of the active site histidine to a tyrosine (residue 191 in Figure 4.1) in CP_{SYN} is also the result of two base pair changes i.e. “CAT” (histidine) in C2 to “TAC” (tyrosine) in CP_{SYN}.

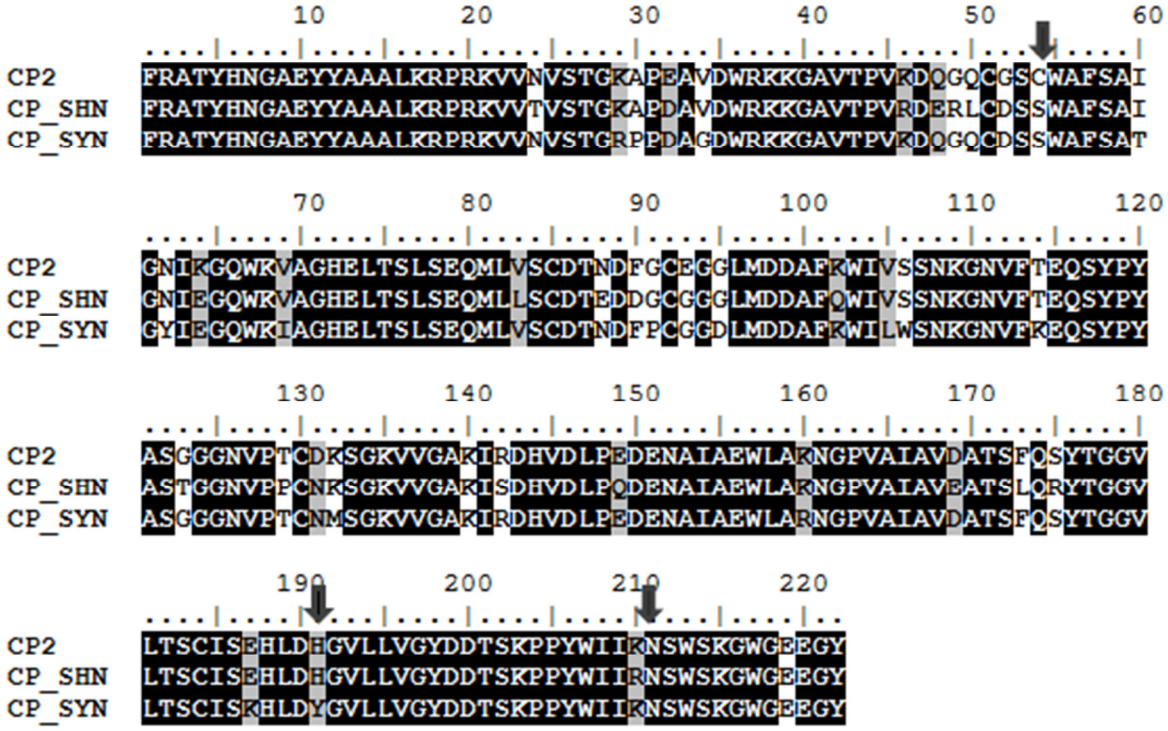


Figure 4.1 Partial amino acid sequence of the two congopain-like CPs selected for expression in *Pichia pastoris*. CP2 (congopain, GenBank® ID: L25130) is included as a reference. Arrows indicate the position of the catalytic triad residues (CHN in CP2). Sequences are named according to their catalytic triad residues. Alignment conducted using ClustalW software (Chenna *et al.*, 2003). Font on a black background (E) indicates identities, white background (E) indicates differences and grey background (E) indicates similarities.

CP_{SHN} and CP_{SYN} were expressed as soluble 28 kDa proteins and secreted into the culture supernatant consistent with expression using the pPic9 expression vector (Figure 4.2 A and B). Given that *P. pastoris* GS115 secretes few native proteins, the supernatant fluid after expression contains few other contaminating proteins (Figure 4.2 A). The identity of the expressed proteins was confirmed by size, 28 kDa which is the same as the control recombinant C2 (Figure 4.2 A and B, lane 3), and the recognition of the expressed proteins by an affinity purified anti-congopain N-terminal peptide antibody (Figure 4.2 B). Separation of CP_{SHN} and CP_{SYN} from culture supernatants by TPP, ion-exchange and molecular exclusion chromatography allowed purification to homogeneity (Figure 4.2 C).

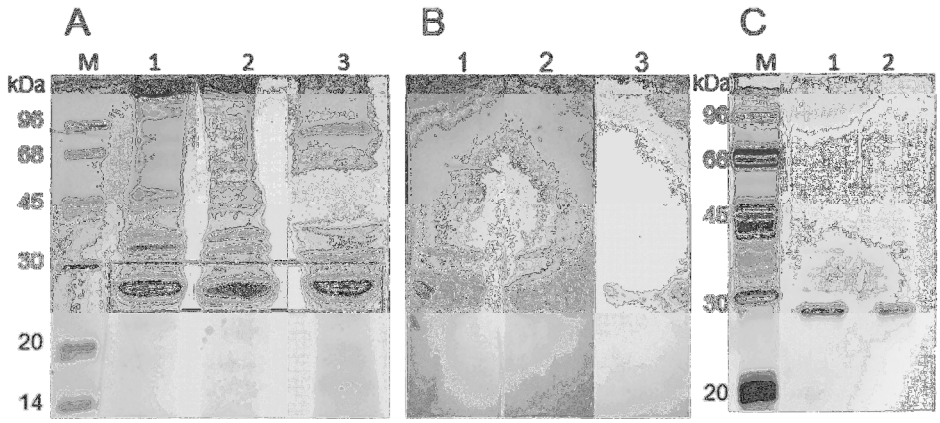


Figure 4.2 Analysis of expression in *Pichia pastoris* (A and B) and purification (C) of CP catalytic domain CP_{SHN} and CP_{SYN} by reducing 12% SDS-PAGE (A and C) and Western blotting (B). Lane M: Bio-Rad low molecular weight marker; lane 1: CP_{SHN}, lane 2: CP_{SYN} and lane 3: C2. (A and B) Supernatant fluids and (C) purified protein were added to loading dye containing 0.7% β-mercaptoethanol and boiled for 2 min prior to loading. The box indicates the bands of interest. (A) Coomassie stained. (B) Primary antibody was affinity purified chicken anti-congopain N-terminal peptide IgY (1 µg/ml). (C) Silver stained.

Table 4.1 shows the purification of the CP variants from *P. pastoris* supernatants after four days of induction. The first TPP purification step resulted in the greatest removal of contaminating proteins as is evident by the decrease in the total protein levels. Anion exchange chromatography also removed contaminating protein, however, it is likely that some of the protein of interest was lost in this step as indicated by the similar specific activities when compared to the TPP step. The final MEC purification step appeared to aid purification of CP_{SHN} to a larger extent that CP_{SYN} as indicated by 6-fold difference in the specific activity values of the two enzymes. The final yield of CP_{SHN} and CP_{SYN} obtained was 33.7 and 5.3% of the total activity in the culture supernatant respectively. However, these figures may not be accurate due to the overestimation of total protein in the initial supernatant as polysaccharides present in the media would have artificially inflated the absorbance reading used to calculate the amount of protein. Therefore, the true yield of both enzymes is likely to be higher than indicated in Table 4.1.

Table 4.1 Purification of recombinant CP_{SYN} and CP_{SHN} from *Pichia pastoris* supernatants after 4 days of induction. Enzyme activity was assayed using Z-Phe-Arg-AMC as a substrate**.**

Step	Total protein (mg)		Total activity (pmol.s^{-1})		Specific activity (pmol.s^{-1}/mg)		Purification (-fold)		Yield (%)	
	CP_{SYN}	CP_{SHN}	CP_{SYN}	CP_{SHN}	CP_{SYN}	CP_{SHN}	CP_{SYN}	CP_{SHN}	CP_{SYN}	CP_{SHN}
Supernatant	125	54.15	7125	11549	57	213	1	1	100	100
TPP[1]	10	15.24	6100	9248	610	606	11	2.8	85.6	80.1
IEC[2]	1.25	12	868	8171	694.4	681	13	3.2	12.2	70.8
MEC[3]	0.5	0.8	379	3887	758	4858	15	23	5.3	33.7

[1]Three phase partitioning, [2]ion exchange chromatography, [3]molecular exclusion chromatography are all further described in Section 4.2.2.

4.3.2 Enzymatic characterisation of CP_{SHN} and CP_{SYN}

The purified proteins were analysed by a gelatin gel for proteolytic activity (Figure 4.3). Interestingly, CP_{SHN} revealed a much higher molecular weight compared to the other two proteins i.e. higher than 68 kDa compared to *ca* 40 kDa. Since gelatin substrate gel analysis is conducted under non-reducing conditions (i.e. no β-mercaptoethanol in the loading buffer and samples not boiled prior to loading), both intra-molecular and inter-molecular disulfide bridges would still be intact. Thus the conformation of the enzyme necessary for catalysis was maintained. Furthermore, due to the limited migration in the gelatin gel (40 and 68 kDa) compared to migration under reducing conditions (Figure 4.2; 28 kDa) it can be speculated that the functional unit for catalysis is a dimer in the case of C2 and CP_{SYN} and a larger multimer for CP_{SHN}. The band corresponding to C2 activity (lane 1) appears larger and more prominent than the band in other lanes despite loading of comparable amounts of protein in the different lanes.

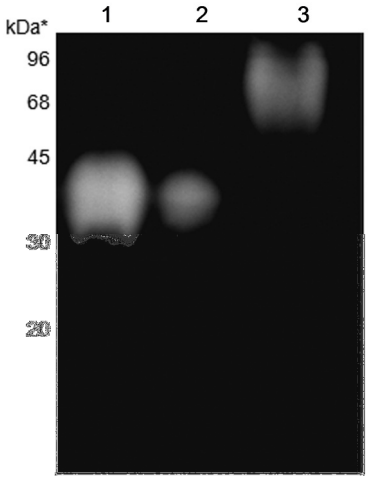


Figure 4.3 Gelatin-containing non-reducing 12% SDS-PAGE to show proteolytic activity of expressed and purified CP_{SHN} and CP_{SYN}. Protein samples were not boiled or reduced and 10 µg of protein was loaded/lane. Lane 1: C2; lane 2: CP_{SYN}; lane 3: CP_{SHN}. *Note that kDa sizes are approximate under non-reducing conditions.

Active site titration was conducted in order to determine the concentration of active enzyme in each expressed variant preparation using E-64 as an active site titrant. For C2 and CP_{SYN}, only 20% of the respective proteins showed enzymatic activity. CP_{SHN}, however, contained 80% active enzyme. Surprisingly, CP_{SHN} and CP_{SYN} showed activity against Z-Phe-Arg-AMC and inhibition by E-64. To our knowledge, this is the first report of such activity shown by cathepsin L-like peptidases.

The kinetic constants for the expressed, purified variant enzymes against the various substrates are shown in Table 4.2. C2 showed the lowest K_m of 6.98 µM and the highest specificity rate constant (k_{cat}/K_m ratio) of 89.22 $s^{-1}mM^{-1}$ against Z-Phe-Arg-AMC. The k_{cat}/K_m values were lower for CP_{SHN} and CP_{SYN} against Z-Phe-Arg-AMC indicating lower hydrolysis of the substrate. C2 showed the lowest K_m value of 8.65 µM for the substrate H-Pro-Phe-Arg-AMC while CP_{SHN} did not cleave the substrate at all.

Table 4.2 Kinetic constants for the hydrolysis of five* synthetic peptide substrates by CP_{SYN} and CP_{SHN} and C2 as a control. Values calculated using the means of triplicate experiments ± SE. K_m was calculated using the Hyper32 software. Assays were carried out in congopain assay buffer [100 mM Bis-Tris pH 6.0, 4 mM EDTA, 0.02% (w/v) NaN_3, 6 mM DTT] with 15 ng of enzyme at 37°C. NR – no reaction after 4 h at 37°C.

Substrate	Enzyme	K_m (µM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)
Z-Phe-Arg-AMC	C2	6.98 ± 0.65	0.62 ± 0.015	89.22
	CP_{SYN}	13.36 ± 1.67	0.29 ± 0.015	21.67
	CP_{SHN}	10.11 ± 1.30	0.17 ± 0.071	17.48
H-Pro-Phe-Arg-AMC	C2	8.65 ± 0.88	1.56 ± 0.34	180.55
	CP_{SYN}	12.05 ± 0.86	0.50 ± 0.051	41.27
	CP_{SHN}	NR	NR	NR
H-Ala-Leu-Lys-AMC	C2	1.32 ± 0.37	1.34 ± 0.042	1008.36
	CP_{SYN}	2.13 ± 0.36	0.51 ± 0.10	239.77
	CP_{SHN}	0.88 ± 0.12	0.27 ± 0.018	309.60
H-Ala-Phe-Lys-AMC	C2	33.18 ± 0.47	0.67 ± 0.031	20.32
	CP_{SYN}	39.35 ± 0.12	0.32 ± 0.036	8.21
	CP_{SHN}	NR	NR	NR
H-Val-Leu-Lys-AMC	C2	2.13 ± 0.19	1.42 ± 0.18	667.54
	CP_{SYN}	2.38 ± 0.45	1.61 ± 0.022	677.18
	CP_{SHN}	2.94 ± 0.34	0.35 ± 0.028	118.88

*No hydrolysis was observed for the substrates Z-Arg-Arg-AMC or Suc-Leu-Tyr-AMC

CP_{SHN} showed the highest affinity for the substrate H-Ala-Leu-Lys-AMC since it had the lowest K_m of 0.88 µM. The k_{cat}/K_m ratios for CP_{SYN} and C2 were similar for H-Val-Leu-Lys-AMC. For the all substrates except H-Ala-Leu-Lys-AMC, C2 had the lowest K_m value and highest specificity rate constant indicating highly efficient hydrolysis. CP_{SHN} was unable to hydrolyse all the tested substrates and did not react at all with H-Pro-Phe-Arg-AMC or H-Ala-Phe-Lys-AMC and showed inefficient hydrolysis of Z-Phe-Arg-AMC. However, CP_{SHN}

did cleave H-Ala-Leu-Lys-AMC and H-Val-Leu-Lys-AMC with low K_m values and high specificity rate constants indicating fairly efficient hydrolysis.

The activity of the CP variants at different pH values was investigated by incubation of the enzymes in constant ionic strength AMT buffers of differing pH before measuring activity against the substrate H-Ala-Leu-Lys-AMC (Figure 4.4). CP_{SHN}, C2 and CP_{SYN} all share a pH optimum of 6.5 regardless of salt concentration (data not shown). CP_{SHN} is active over a narrower pH range than C2 and CP_{SYN}.

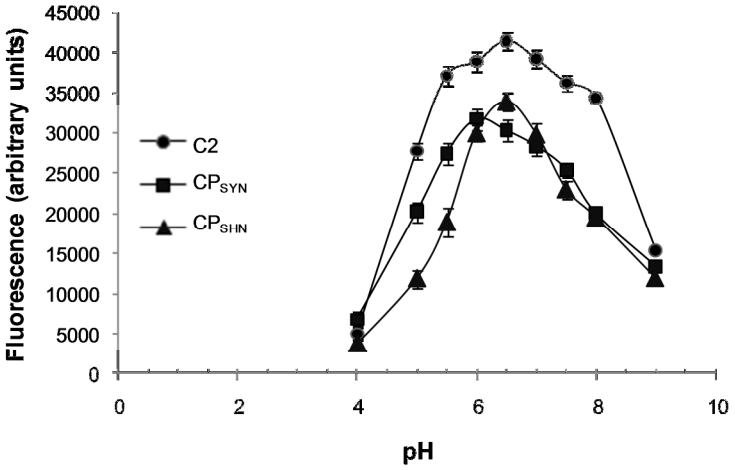


Figure 4.4 pH activity profiles of C2, CP_{SYN} and CP_{SHN} over the pH range of 4.0 to 9.0 using H-Ala-Leu-Lys-AMC. Enzymes (1.5 ng) were assayed for activity after incubation in AMT (100 mM Na-acetate, 200 mM Tris-Cl, 100 mM MES, 6 mM DTT, 4 mM Na_2EDTA) buffers of different pH at 37°C.

The sensitivity of CP_{SHN} and CP_{SYN} variants to the peptidase inhibitors E-64, leupeptin, AEBSF and pepstatin A was tested by incubating the enzyme with inhibitor and subsequently measuring the residual activity against H-Ala-Leu-Lys-AMC. Like C2, both the variants were sensitive to the classical cysteine peptidase inhibitor E-64, since addition of this inhibitor abolished almost all enzyme activity (96-98% inhibition). CP_{SHN} and CP_{SYN} were similarly inhibited by leupeptin (an inhibitor of both serine and cysteine peptidases) (89-96% inhibition). For CP_{SYN} and C2 no significant inhibition was noted using AEBSF (a serine peptidase inhibitor) or pepstatin A (an inhibitor of aspartic peptidases), although CP_{SHN} was slightly inhibited (11 and 9% inhibition respectively).

Sensitivity of the expressed peptidases to diazomethyl ketone inhibitors Z-Phe-Phe-DMK and Z-Phe-Ala-DMK was investigated by measuring the association constant (Table 4.3). CP_{SHN} was most sensitive to both DMK inhibitors with a K_{ass} value an order of magnitude higher than that of the other enzymes for the inhibitor Z-Phe-Ala-DMK. CP_{SYN} was the least sensitive to inhibition by the DMK inhibitors.

Table 4.3 Effect of diazomethyl-ketone inhibitors on activity of C2, CP_{SYN} and CP_{SHN}. Enzyme (10 mM) was incubated in congopain assay buffer at 37°C for 5 min. Inhibitor (10 mM) and substrate (H-Ala-Leu-Lys-AMC, 20 µM**)** were added and initial rates measured allowing calculation of the second order rate constant.

Inhibitor	Z-Phe-Phe-CHN_2		Z-Phe-Ala-CHN_2	
	k_{ass} (× 10^{-5} /Ms)	$t_{½}$ (s)	k_{ass} (× 10^{-5}/Ms)	$t_{½}$ (s)
C2	4.5 ± 0.2	483	1.5 ± 0.5	279
CP_{SYN}	1.9 ± 0.2	1137	0.7 ± 0.3	178
CP_{SHN}	16.5 ± 3.5	44	10.0 ± 1.4	115

4.4 DISCUSSION

In the present study two congopain-like variant peptidases were expressed in a yeast expression system to investigate whether they differ from congopain enzymatically and biochemically. Although expression of active full length cruzain, the major cysteine peptidase of *T. cruzi* has been successfully conducted in *E. coli* (Eakin *et al.*, 1993), it has been found that expression of congopain in *E. coli* does not yield active enzyme even after refolding of the expressed protein (Boulangé *et al.*, 2001). Functional expression of the catalytic domain of congopain was finally achieved using a baculovirus expression system although with mediocre yield (1 mg/l culture) (Boulangé *et al.*, 2001).

Recently, high level expression of functional congopain, both truncated catalytic domain and full length protein, has been conducted in the methylotrophic yeast *P. pastoris* (Boulangé et al., 2011). This is consistent with the successful expression of many heterologous proteins in this system, the most pertinent being the expression of rhodesain, the major cysteine peptidase of *T. b. rhodesiense* (Caffrey *et al.*, 2001), multiple isoforms of a cathepsin-B like peptidase from *Trichobilharzia regenti,* an avian parasite (Dvořák *et al.*, 2005) and cathepsin B of *T. congolense* (Mendoza-Palomares *et al.*, 2008). The main advantage of the *P. pastoris* expression system over bacterial expression systems is that proteins are eukaryotically processed i.e. proteolytic processing, folding, disulfide bond formation, glycosylation and other post-translation modifications are all performed, thereby eliminating the need for refolding after expression (Aloulou *et al.*, 2006). Also, most *P. pastoris* expression vectors contain the *Saccharomyces cerevisiae* α-factor signal-peptide that has been shown to be an effective secretion signal sequence. This extracellular expression, in addition to the fact that *P. pastoris* secretes very few proteins, greatly simplifies the purification of expressed proteins.

CP_{SHN} and CP_{SYN} were successfully expressed as shown by the 28 kDa bands on reducing SDS-PAGE. A gelatin gel showed that CP_{SHN} migrated at 68 kDa under non-reducing conditions as opposed 45 kDa that was the case for CP_{SYN} and C2. A possible explanation for this is that CP_{SHN} might be active as a stable multimer. Interestingly, it has previously been shown that congopain is active as a dimer *in vivo* although the exact residues involved in dimerisation are still under investigation (E. Authié and A. Boulangé, personal communication).

Activity of the two expressed enzymes, in comparison to C2 was tested against five substrates. Substrates with a hydrophobic amino acid residue in P_1 were not cleaved at all by either of the variants or C2. The K_m value for C2 against Z-Phe-Arg-AMC obtained in the present study (6.98 µM) compares favourably to those obtained in previous studies for native congopain of 1.5 µM (Chagas *et al.*, 1997), 4.4 µM (Mbawa *et al.*, 1992) and 7.3 µM (Authié *et al.*, 1992).

The K_m values for CP_{SYN} for the hydrolysis of Z-Phe-Arg-AMC and H-Ala-Leu-Lys-AMC suggest that CP_{SYN} needs P_3 to be occupied by a small, neutral amino acid rather than the bulky cyclic amino acid proline (H-Pro-Phe-Arg-AMC). It is evident that the two expressed enzymes, CP_{SHN} and CP_{SYN} have substrate specificities slightly different to that of C2. However, the two congopain-like CPs investigated in this study all prefer substrates with a hydrophobic amino acid in P_2 and amino acids with a polar side chain in P_1 which is consistent for what was shown for congopain (Lecaille *et al.*, 2001a). The multiple cathepsin L-like enzymes found in *T. cruzi* and *L. major* also display different substrate preferences leading to the idea that the enzymes perform divergent functional roles *in vivo* (Mottram et al., 1997; Lima et al., 2001).

An interesting finding from the present study is that CP_{SHN} in which the catalytic cysteine has been replaced by a serine is sensitive to the class-specific cysteine peptidase inhibitor, E-64, and insensitive to the serine peptidase inhibitor AEBSF. The serine repeat antigens (SERA) of the *Plasmodium* species are a good example of other cysteine peptidases with unusual catalytic triads (Aoki *et al.*, 2002). SERA5 is of particular interest since the enzyme possesses a papain-like central domain with the catalytic cysteine replaced by a serine (Hodder *et al.*, 2003). Expression and characterisation of SERA 5 showed that the enzyme behaved as a serine peptidase rather than a cysteine peptidase, since it was inhibited by the classical serine peptidase inhibitor, 3,4-dichloroisocoumarin (Hodder *et al.*, 2003), which is not the case with the serine mutants expressed in the

present study (CP_{SHN} and CP_{SYN}) which were not inhibited by AEBSF. However, there is evidence that E-64 may bind to and subsequently inhibit the action of serine peptidases via the S_1 site (Sreedharan et al., 1996). Several of the genes coding for a novel family of cathepsin B-like peptidases expressed in bloodstream form *T. congolense* code for peptidases where the nucleophilic cysteine is replaced by a cysteine, but these peptidases did not show activity (Mendoza-Palomares *et al.*, 2008). It has been suggested that a cysteine-to-serine substitution in peptidases of haemoparasites could be a way to moderate the oxidation of the essential catalytic cysteine in the oxygen-rich environment of the blood (Atkinson *et al.*, 2009). Another example of an unusual catalytic triad is found in the poliovirus peptidase 3C. The poliovirus peptidase 3C is structurally similar to chymotrypsin but the catalytic serine residue has been replaced by a cysteine residue. This enzyme functions as a cysteine peptidase in terms of substrate specificity and sensitivity to inhibitors (Sarkany and Polgar, 2003).

A pH profile of the CP_{SHN} and CP_{SYN} was compared with that of C2. Both variants showed an optimum of activity close to physiological pH (6.5) similar to the earlier finding for congopain (Mbawa et al., 1992; Chagas et al., 1997). However, the activity is minimal at pH 4.0 and below, a paradoxical occurrence for a lysosomal enzyme.

In this study, two variant congopain-like enzymes were cloned and expressed in *P. pastoris*. Substrate assays showed that the two variant CPs had different substrate preferences compared to C2, as evident from the catalytic constants. Also of interest is that the expressed variants of congopain showed an optimum for catalytic activity at pH 6.5. These results indicate that it is possible that these variants are expressed in the parasite and may play distinct roles *in vivo.* However, more work on the function and exact mechanism of catalysis of these variant congopain–like enzymes needs to be conducted.

CHAPTER 5: FUNCTIONAL EXPRESSION OF TWO ALANYL AMINOPEPTIDASES OF *TRYPANOSOMA CONGOLENSE*

D. Pillay[1*], A. Boulangé[2], T. Baltz[3] and T.H.T. Coezter[1]

[1] School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal (Pietermaritzburg campus), Private Bag X01, Scottsville 3209, South Africa
[2] Trypanosomoses, UMR-IRD/CIRAD 017, CIRAD-Emvt, 34398 Montpellier, France
[3] Laboratoire de Microbiologie Cellulaire et Moléculaire et Pathogénicité, UMR-5234 CNRS, Bordeaux, France
[*] Email: 202511575@ukzn.ac.za

Trypanosoma congolense **is a haemoprotozoan parasite that causes African animal trypanosomosis, a wasting disease of cattle and small ruminants. Current control methods are unsatisfactory and no conventional vaccine has been developed yet. An anti-disease vaccine approach to *T. congolense* has been proposed, requiring the identification of factors within the parasite which cause disease. Immunoprecipitation of *T. congolense* antigens using sera from infected cattle has allowed the identification of several potential pathogenic factors, including two M1 type aminopeptidases. Both aminopeptidases were cloned and expressed in *E. coli*. As the aminopeptidases expressed as insoluble inclusion bodies it has been necessary to develop a method for solubilisation and subsequent refolding of the aminopeptidases to restore conformation and activity. The active aminopeptidases showed a substrate preference for H-Ala-AMC, an optimum pH of 8.0, localisation in the cytoplasm, and inhibition by puromycin, all characteristics of cytosolic M1 type aminopeptidases. The two aminopeptidases are not developmentally regulated, as was shown by their presence in procyclic, metacyclic and bloodstream form parasites. Down-regulation of both aminopeptidases by RNAi resulted in a slightly reduced growth rate in procyclic parasites *in vitro*. Immunisation of BALB/c mice with the aminopeptidases did not provide protection when challenged with *T. congolense*.**

Trypanosoma congolense is a protozoan parasite causing the debilitating wasting disease, African animal trypanosomosis (AAT) in cattle and small ruminants. *T. congolense* is transmitted by the tsetse fly which infests 10 million km^2 in sub-Saharan Africa (Krafsur, 2009). AAT is of great economic concern since the disease results in losses from USD $ 1 - 5 billion per annum in Africa (Kristjanson *et al*., 1999). Current control measures for AAT include the use of trypanocides, control of the tsetse vector and the use of trypanotolerance cattle, which have an innate genetic resistance to the disease (Torr et al., 2005; d'Ieteren and Kimani, 2007). However, these control measures are insufficient to curb the spread of AAT. Vaccination would be a viable alternate control method for AAT, but antigenic variation, the immune evasion mechanism used by Salivarian trypanosomes, has greatly hindered the development of a conventional vaccine (Morrison et al., 2009; Baral, 2010). For these reasons, an anti-disease approach has been put forward for the development of an AAT vaccine (Authié et al., 2001; Antoine-Moussiaux et al., 2009). This requires the identification of all the factors of the parasite which contribute to disease in the host animal, i.e. pathogenic factors.

The peptidases of protozoan parasites have drawn attention as potential drug targets, pathogenic factors and vaccine candidates (Sajid and McKerrow, 2002; McKerrow et al., 2006; Caffrey and Steverding, 2009). The cathepsin L-like (Lalmanach et al., 2002; Pillay et al., 2010; Boulangé et al., 2011) and cathepsin B-like (Mendoza-Palomares *et al.*, 2008) peptidases of *T. congolense* have been studied. Interestingly, both these groups of peptidases are encoded by multiple genes with slight variations in sequences resulting in peptidases with varying catalytic sites. Furthermore, both cathepsin B and L-like peptidases are expressed only in the bloodstream form of the parasites, making them attractive targets for therapeutic drugs.

Aminopeptidases (APs) are metal ion dependent peptidases that catalyse the removal of single N-terminal amino acid residues from peptides thus playing a role in protein maturation and degradation (Taylor, 1993). In *Plasmodium*, APs have a critical role in degradation of haemoglobin to produce certain amino acids required for protein synthesis which the parasite cannot produce itself (Skinner-Adams *et al.*, 2009). Consequently, APs have been shown to have potential application as a drug target in *P. falciparum* and *T. brucei* (Knowles, 1993; Flipo et al., 2007) and a vaccine candidate in *F. hepatica* (Piacenza et al., 1999; Acosta et al., 2007). In spite of this, none of the APs of the pathogenic African trypanosomes have yet been studied.

Alanyl or Family M1 type APs are characterised by a metal binding “HEXXH” motif (Figure 5.1) whereby the two histidine residues are responsible for the binding of a zinc ion that serves as a catalytic co-factor, common to all metal binding APs (Clan MA) (Rawlings and Barrett, 1993). A second substrate binding motif, “GXMEN” is specific to M1 type APs and the glutamic acid residue is implicated in the mechanism of catalysis since mutation of this residue has been shown to decrease or abolish AP activity (Vazeux et al., 1998; Laustsen et al., 2001).

Using immuno-affinity isolation, two M1 type APs of *T. congolense* were identified as being potential pathogenic factors (see Chapter 2). In the present study, these APs were subsequently cloned, expressed, purified, characterised enzymatically, and localised in the parasite life cycle stages using immunofluorescence. Immunisation and infection experiments were conducted to determine if there was any level of protection conferred by immunisation with the APs. RNAi experiments were conducted to determine whether the

APs play an essential role in the parasite. This information could contribute to further studies of these M1 type APs as either vaccine candidates or drug targets.

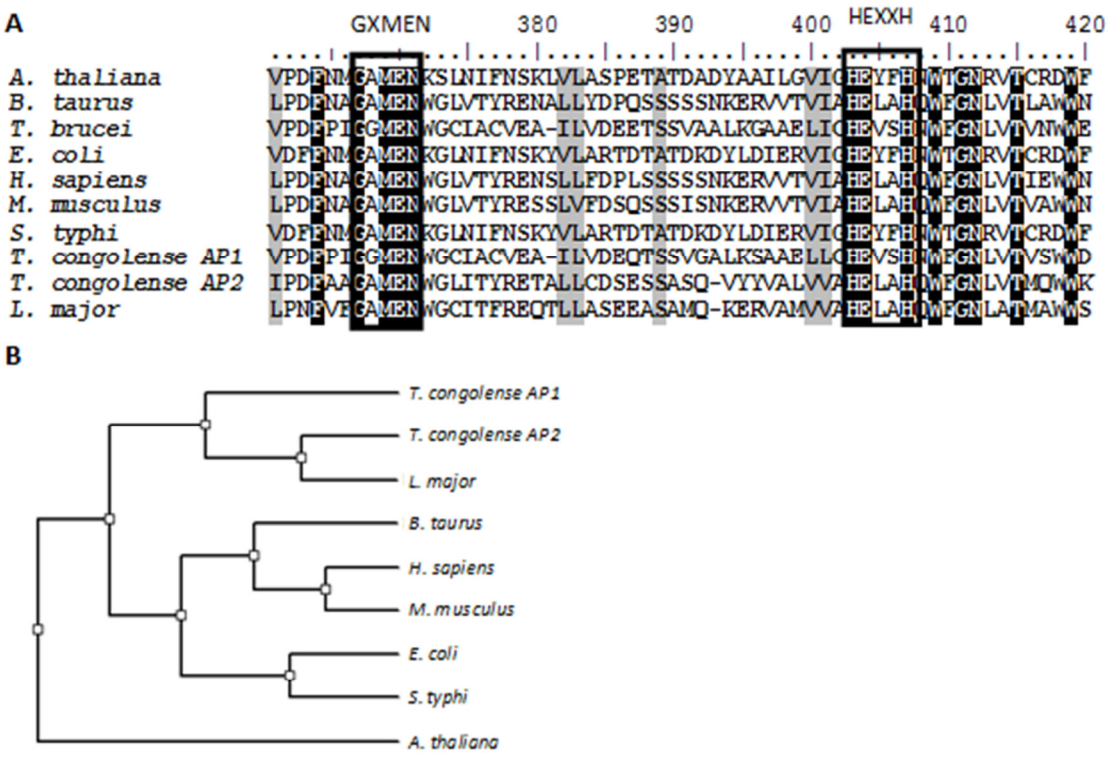


Figure 5.1 The phylogenetic relationships between the two *Trypanosoma congolense* Family M1 aminopeptidases and other Family M1 aminopeptidases. **(**A) Multiple sequence alignment of a part of the catalytic domain of Family M1 aminopeptidases. The HEXXH zinc-binding motif and the GXMEN motif, specific to the M1 Family, are boxed. Sequences were obtained from GenBank® (http://www.ncbi.nlm.nih.gov/genbank/, accessed 5-8-2009): *Arabidopsis thaliana* (AAX59049.1), *Bos taurus* (NP001068612.1), *Trypanosoma brucei* (AAS67871.1), *Escherichia coli* (89107782), *Homo sapiens* (P5144), *Mus musculus* (31077182), *Salmonella typhi* (CAD08183.1), *Leishmania major* (CA5045941). (B) Phylogenetic tree indicating the relationships between full protein sequences of several M1 aminopeptidases. Tree created using the Prot Distance-Neighbour phylogenetic tree accessory application in BioEdit (Hall, 1999).

5.1 MATERIALS

General molecular biology: NotI, NdeI, O'RangeRuler® DNA Ladder 200 bp and the GeneJet® Plasmid Miniprep Kit were obtained from Fermentas (Vilnius, Lithuania). *Taq* polymerase was obtained from Solis Biodyne (Tartu, Estonia). The Wizard® PCR Preps DNA Purification System, the Wizard® Plus SV Minipreps DNA Purification System and thrombin were purchased from Promega (Madison, WI, USA). The illustra GFX® PCR DNA and Gel Band Purification Kit was purchased from GE Healthcare (Buckinghamshire, UK). For sequencing, the Big Dye® Terminator v3.1 Cycle Sequencing Kit was obtained from Applied Biosystems (Foster City, CA, USA).

***E. coli* cells:** Competent *E. coli* cells (NEB 5-alpha K12 and One Shot® Top10F') were purchased from New England Biolabs (Ipswitch, MA, USA) and Invitrogen (Carlsbad, CA, USA) respectively. Both *E. coli* strains used allow blue/white screening for transformants in the presence of X-gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) by α-complementation of the β-galactosidase gene.

Protein purification: The BugBuster® Protein Extraction Reagent was obtained from Novagen (Darmstadt, Germany). Protein Refolding Kit® and Restore® Western Blot stripping buffer were bought from Pierce (Rockford, IL, USA). The HisTrap® HP columns, the Q-sepharose® HP columns, the ÄKTApurifier® and the enhanced chemoluminescence (ECL) substrate kit were obtained from GE Healthcare (Buckinghamshire, UK). Complete Protease Inhibitor Cocktail Tablets® was from Roche (Mannheim, Germany). Bradford Protein Assay Reagent was obtained from Bio-Rad (Hercules, CA, USA).

Peptide substrates and inhibitors: The peptide substrates H-Leu-AMC, H-Gly-AMC, H-Arg-AMC, H-Tyr-AMC, H-Ala-AMC, and H-Met-AMC were obtained from Bachem (King of Prussia, PA, USA). *Trans*-epoxysuccinyl-L-leucyl-amido(4-guanidino)butane (E-64), leupeptin, pepstatin and 4-2-(Aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF) were purchased from Sigma (Munich, Germany). Fluorescence assays were measured using a FLUOR Star Optima spectrophotometer from BMG Labtech (Offenburg, Germany).

Antibody preparation: The peptides chosen for antibody production were synthesised by Auspep (Parkville, Australia). Sephadex® G-10 and Sephadex® G-25 resins, Freund's complete and incomplete adjuvants and rabbit anti-IgY coupled to HRPO were purchased from Sigma. Nunc 96 well plates were obtained from ThermoFisher Scientific (Roskilde, Denmark). Aminolink® and Sulfolink® matrices were purchased from Pierce.

Mouse immunisation: BALB/c mice were obtained from the National Health Laboratory Services (Pretoria, South Africa) and NOD/SCID (NOD.Cg-Prkdescid Il2rgtm1 Wjl/Szj) were bred at the University of Bordeaux 2 (Bordeaux, France) in specific pathogen-free conditions and used for experiments from five to six weeks of age. Protocols using mice were approved by the University of KwaZulu-Natal animal ethics committee (Reference 036/09/Animal) and the University of Bordeaux 2 animal care and use committee and the Commission de Genie Génetique (Direction Generale de la Recherche et de l'Innovation). The haematocrit centrifuge was purchased from Heraeus (Frankfurt, Germany).

Trypanosome culture: Culture flasks (25 cm^2) and 24-well plates were obtained from Corning (NY, USA). Amaxa Nucelofaction® transformation system and the Amaxa Basic Parasite Nucleofactor® 2 buffer were obtained from Lonza (Levallois-Perret, France). MEM Base powder, foetal calf serum, HEPES buffer, hypoxanthine, haemin, glutamine, proline, bathocuproin, thymidine, goat serum and adenine were obtained from Sigma.

Immunofluorescence: Chicken anti-BiP69 IgY was a gift from S. Hadebe (UKZN, PMB, SA). Monoclonal mouse anti-tubulin antibodies were a gift from K. Gull (Oxford, UK). Secondary antibodies Alexa Fluor goat anti-mouse IgG 488, Alexa Fluor goat anti-chicken IgG 488 and Texas Red® goat anti-mouse IgG and *SlowFade*® Gold anti-fade reagent were obtained from Invitrogen. The MetaMorph® software was provided by Molecular Devices (CA, USA).

5.2 METHODS

5.2.1 Cloning, expression, purification and refolding of APs

Genomic DNA was isolated from a pellet of *T. congolense* (strain IL3000) according to the method of Medina-Acosta and Cross (1993). Briefly, the trypanosome pellet was washed with 1 ml PBS and centrifuged (10 min, 2000 *g,* RT). The pellet was resuspended in 150 µl of TELT buffer [50 mM Tris-Cl, pH 8.0, 62.5 mM EDTA, pH 8.0, 2.5 M LiCl, 4% (v/v) Triton X-100] and incubated at RT for 5 min. A phenol-chloroform extraction was performed. The top phase was recovered and 300 µl of 100% (v/v) ethanol added and mixed by inversion for 5 min. The sample was centrifuged (10 min, 10 000 *g*, RT) and the supernatant discarded. The pellet was washed with 1 ml of 70% (v/v) ethanol and the pellet allowed to dry at 37°C. The pellet was resuspended in sterile distilled water with 20 µg/ml RNAse and incubated at 37°C for 45 min.

Primers were designed to amplify the full length open reading frames for AP1 (Forward: 5' AAA ACA TAT GTC AGC CGC CGT TGC AAC AGT GAA GC 3'; Reverse: 5' AAA GCG GCC GCC TAC CGT GGA AAA AAT AAG TA 3 ') and AP2 (Forward: 5' AAA ACA TAT GCC CGG AAC TCG TGA TGT T 3'; Reverse: 5' AAA GCG GCC GCT TAT TCT GCG TTG CTG GAG AG 3'). These specific primers were used under the following conditions : 10 cycles at 95°C for 20 s, 55°C for 30 s, 68°C for 2 min followed by 25 cycles at 94°C for 20 s, 55°C for 30 s, 68°C for 2 min increasing by 10 s every cycle, to reduce non-specific amplification. The primers include restriction sites (underlined) for NdeI (Forward primer) and NotI (Reverse primer) to facilitate cloning into the expression vector. The PCR

products were purified from an agarose gel using the illustra GFX® PCR DNA and Gel Band Purification Kit (GE Healthcare). All agarose gels used contained 1% (w/v) agarose unless otherwise specified. The PCR products were subjected to a restriction digest with the enzymes NdeI and NotI (Fermentas, Vilnius, Lithuania) as per the manufacturer's instructions. The restricted PCR products were again subjected to a gel purification and subsequently ligated into a similarly restricted expression vector (pET28a) allowing fusion with an N-terminal 6 x His-tag. The AP genes were also cloned into the pGEX-4T-1 and pET32a vectors. Furthermore, expression was initially attempted in *Pichia pastoris* using the pPicZαA28 and pPic9 vectors (see Section 4.2.1 for method).

Competent cells (*E. coli* JM109) were transformed by heat shock according to the manufacturer's instructions and grown on LB containing 33 µg/ml kanamycin for pET28a for selection of transformants. A small number of transformed colonies were subjected to colony PCR using primers specific for the vector (pET Forward: 5' TAA TAC GAC TCA CTA TAG GG 3'; pET Reverse: 5' GCT AGT TAT TGC TCA GCG G 3') to select recombinant clones. Recombinant colonies were grown overnight in LB with 33 µg/ml kanamycin at 37°C. Minipreps were performed using 5 -10 ml of overnight culture with the Wizard® Plus SV Minipreps DNA Purification System. Cloned fragments were checked for the absence of mutations by sequencing (Big Dye® Terminator v3.1 Cycle Sequencing Kit). Internal primers were used to obtain the entire sequence (approximately 2.5 kb) (AP1 Forward : 5' GCA AGT ATC CGT TGC 3'; AP2 Forward: 5' GCC TGA ATT ATG TGC 3'). Recombinant plasmids were transformed into *E. coli* BL21 (DE3) for expression.

For expression, an overnight culture of LB with kanamycin was prepared. The overnight culture was diluted 1/50 in fresh medium and allowed to grow for approximately 2 h at 37°C until the OD_{600} was between 0.6 and 0.8. Expression was induced by addition of 0.4 mM IPTG to the culture. Cultures were centrifuged (2000 *g*, 10 min, RT) to collect cells which were either purified immediately or stored at -20°C until purification.

Cell pellets were resuspended in BugBuster® Protein Extraction Reagent (1/10 culture volume) and sonicated (5 x 10 s pulses). The cell lysate was centrifuged (10 000 *g*, 30 min, 4°C) and the supernatant discarded. The inclusion bodies were washed three times in 1/10 volume of extraction buffer [8 M urea, 2.5% (v/v) Triton X-100, 0.5 M NaCl] and centrifuged (10 000 *g*, 30 min, 4°C). The pellet was resuspended in 8 M guanidine (dissolved in 100 mM Na-phosphate, pH 8.0, 10 mM DTT) using gentle agitation at RT for

at least 12 h. Undissolved debris was removed by centrifugation (10 000 *g*, 30 min, 4°C) and the supernatant diluted to approximately 0.1 mg/ml for refolding.

Refolding buffer was optimised for guanidine (0.4 – 1.5 M) and L-arginine (0 – 0.8 M) concentrations as well as ratio of reduced to oxidised glutathione (1:1 to 2:0.4 mM) using the Pierce® Protein Refolding Kit. The optimised refolding buffers [55 mM Tris-Cl buffer, 21 mM NaCl, 0.88 mM KCl, pH 8.2, 0.4 M guanidine, 2 mM reduced glutathione, 0.4 mM (0.2 mM for AP2) oxidised glutathione, 0.4 M (omitted for AP2) L-arginine] was used to slowly dilute the crude inclusion body preparation and incubated at 4°C for 24 h. Refolded samples were centrifuged (10 000 *g*, 30 min, 4°C) to removed precipitated debris and dialysed against several changes of 20 mM Tris-Cl buffer, pH 8.0 at 4°C over 24 h.

Refolded protein was partially purified on a 1 ml HisTrap® HP column using the ÄKTApurifier® and a gradient of 50 mM to 1 M imidazole in column buffer (20 mM Tris-Cl, 0.5 M NaCl, pH 8.0). Protein-containing fractions (identified by OD_{280}, SDS-PAGE and activity assays) were pooled and diluted 1/3 in distilled water to decrease the final NaCl concentration, and purified to homogeneity using a Q-Sepharose® HP column using an elution gradient from 160 mM to 1 M NaCl on the ÄKTApurifier®. Protein-containing fractions (0.5 ml each) were pooled and stored in 20% (v/v) glycerol. Removal of the His-tag by thrombin cleavage was optimised to allow complete removal with minimal degradation of the protein as confirmed by western blotting. Following this, optimal conditions (1 unit, 4 h, RT) were routinely used and thrombin was subsequently inhibited using AEBSF (100 µg/ml).

5.2.2 Enzymatic characterisation of AP1 and AP2

Enzyme kinetics for the two recombinant, purified APs were determined for the following substrates: H-Leu-AMC, H-Gly-AMC, H-Arg-AMC, H-Tyr-AMC, H-Ala-AMC, H-Met-AMC. Briefly, enzyme [1.5 ng diluted in 0.1% (v/v) Brij-35] was incubated in assay buffer (50 mM Tris-Cl buffer, pH 8.0, 50mM $MnCl_2$) at 37°C for 10 min. Substrate stocks were diluted to 100, 50, 40, 30, 20, 15, 10, 7.5, 5, 2.5 mM (final concentrations) and 25 µl added to 75 µl of enzyme-buffer mix. Fluorescence was measured (excitation 360 nm; absorbance 460 nm) for thirty cycles using the FLUOR Star Optima spectrophotometer from BMG Labtech (Offenburg, Germany). Michaelis-Menten plots were used to determine the kinetic constants, K_m and V_{max} using the software Hyper32® (1991-1992, J.S. Easterby, University of Liverpool, UK). The specificity constant, k_{cat} was calculated from the formula

k_{cat} = V_{max}/E_o, where E_o is the concentration of active enzyme [assumed at 100% activity since no active site titrant is currently available (Morty and Morehead, 2002)].

Metal cation dependency of the recombinant APs was investigated by supplementing the assay buffer with two different concentrations (1 mM and 10 mM) of various metal cations ($CaCl_2$, $CoCl_2$, $CuCl_2$, $FeCl_2$, $NiCl_2$, $MgCl_2$, $MnCl_2$, and $ZnCl_2$) prior to incubation with enzyme (1.5 ng) for 30 min and analysis of cleavage of H-Ala-AMC.

In order to determine the pH optimum for each AP, a pH profile was generated. Constant ionic strength AMT buffers (100 mM Na-acetate, 200 mM Tris-Cl, 100 mM MES, 6 mM DTT, 4 mM Na_2EDTA) were prepared from pH 4.0 to pH 9.0 in 1.0 increments by titration with NaOH or HCl (Ellis and Morrison, 1982). Enzymes (1.5 ng) diluted in 0.1% (v/v) Brij-35 were incubated in each buffer of a different pH for 5 min at 37˚C. The substrate H-Ala-AMC (20 µM, 25 µl) was added and fluorescence measured (excitation 360 nm; absorbance 460 nm).

The pH stability assay was conducted in a similar way with the following exceptions: the enzyme was incubated with the buffer for 2 h at 37°C; 5 µl of the enzyme-buffer mix was subsequently aliquoted into 95 µl of the optimal pH buffer (pH 8.0 for AP1, pH 7.5 for AP2) prior to addition of substrate (H-Ala-AMC, 20 µM) and measurement of fluorescence. Each experiment was conducted in triplicate and kinetic constants were calculated from the means of three experiments.

Inhibition of the recombinant APs was investigated using the following compounds: AEBSF (1 mM); amastatin (0.05, 0.5, 1.25 µM); bestatin (2.5, 10, 20, 50 nM); DTT (0.5 mM); E-64 (10 µM); reduced glutathione (1 mM); EDTA (0.5 M); leupeptin (10 µM); pepstatin A (10 µM); 1,10 phenanthroline (0.1 mM) and puromycin (1 mM). To initially determine whether any inhibition occurred, enzyme [1.5 ng in 0.1% (v/v) Brij-35 in 25 µl] was mixed with assay buffer (50 µl) containing inhibitor (at various concentrations as listed above) and incubated at 37°C for 15 min before addition of the substrate (H-Ala-AMC, 20 µM). The inhibition constant K_i for non-tight binding, reversible inhibitors was determined as described by Salvesen and Nagase (2001). Firstly, the uninhibited reaction velocity (v_o) of each AP hydrolysing H-Ala-AMC was determined. Following this, a 20-fold molar excess of inhibitor over enzyme was added in less than 5% of the total assay volume, and the inhibited reaction velocity (v_i) calculated. The apparent inhibition constant $K_{i(app)}$ was calculated from the following equation: $V_o / V_i = 1 + [I] / (K_{i(app)})$; where

[*I*] denotes the inhibitor concentration. True K_i could subsequently be derived from: $K_i = (K_{i(app)}) / 1 + ([S]/K_m)$.

5.2.3 Anti-AP1 and AP2 peptide antibody preparation

A 20-mer peptide was selected from each aminopeptidase sequence for raising anti-peptide antibodies using the Predict7 software (Cármenes *et al*., 1989). Peptides were selected on the basis of high hydrophilicity, surface probability, flexibility and antigenicity. Peptides were synthesised to greater than 70% purity as determined by HPLC (Auspep, Australia). The peptides were prepared for immunisation by coupling to rabbit albumin via either glutaraldehyde for the N-terminus of the AP1 peptide or *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester (MBS) for the C-terminus of the AP2 peptide.

Coupling using glutaraldehyde was done according to Briand *et al*. (1985). Peptide (4 mg) was dissolved in 50 µl DMSO, 450 µl PBS (pH 7.4) added before mixing with rabbit albumin (1:40 ratio of peptide : rabbit albumin), dissolved in 500 µl PBS, and finally 2% (v/v) glutaraldehyde (1 ml) was added drop-wise with stirring and mixed gently for 2 h at 4°C. The coupling reaction was stopped by the addition of $NaBH_4$ (10 mg/ml) and stirred for 1 h further at 4°C. The solution was subsequently dialysed against three changes of PBS containing 0.02% (w/v) NaN_3.

For coupling via MBS, the peptide (4 mg) was dissolved in 100 µl DMSO and added to 400 µl reducing buffer [100 mM Tris-Cl buffer, 10 mM EDTA, 0.02% (w/v) NaN_3, pH 8.0]. DTT (10 mM) was added to make a final volume of 2 ml and the solution incubated for 1.5 h at 37°C. The reduced peptide was subsequently separated from excess MBS and uncoupled peptide on a Sephadex® G-10 column (15 x 110 mm, 10 ml/h). The eluate was collected in 500 µl fractions and 10 µl of each fraction was tested with Ellman's reagent [10 mM Ellman's reagent, 100 mM Tris-Cl buffer, 10 mM EDTA, 0.1% (w/v) SDS, pH 8.0]. Yellow fractions containing the reduced peptide were pooled. Rabbit albumin and MBS (dissolved in DMF) were mixed in a 1:40 molar ratio at RT for 30 min with agitation. The MBS-activated rabbit albumin was purified on a Sephadex® G-25 column (15 x 130 mm, 10 ml/h). The reduced peptide was subsequently combined with the activated rabbit albumin for 3 h at RT with agitation.

For each peptide, two chickens were immunised intramuscularaly with 200 µg of rabbit albumin-coupled peptide emulsified with Freund's complete adjuvant. Booster injections, using Freund's incomplete adjuvant were performed at weeks 2, 4 and 6 after the initial injection. Eggs were collected daily for 12 weeks and stored at 4°C until IgY isolation was

conducted. IgY was purified from all the eggs collected from each chicken every week and the IgY isolated from eggs collected prior to the first week of immunisation constituted a pre-immune control. IgY isolation from egg yolks was conducted as per the method of Polson *et al.* (1985). Briefly, egg yolks were separated from egg whites and the yolk sac discarded. Two volumes of IgY buffer [100 mM Na-phosphate buffer, pH 7.6, 0.02% (w/v) NaN_3] was added to the yolk and mixed before addition of 3.5% (w/v) PEG 6000. Upon complete dissolution of the PEG, the sample was centrifuged (4 420 *g*, 20 min, RT) and the supernatant filtered through absorbent cotton wool. The PEG 6 000 concentration of the clear filtrate was increased by 8.5% to 12% (w/v) final concentration and the preparation centrifuged (12 000 *g*, 10 min, RT). The pellet was resuspended in IgY buffer in a volume equal to the original yolk volume and 12% (w/v) PEG 6 000 added and mixed. The sample was centrifuged (12 000 *g*, 10 min, RT) and the pellet resuspended in IgY buffer at 1/6 of the original yolk volume. IgY concentration was determined by measuring the absorbance at 280 nm using the extinction coefficient $E^{1mg/m} = 1.25$ at 280 nm (Goldring *et al.*, 2005). Isolated IgY antibodies were tested by ELISA to determine their titre. Nunc 96 well plates were coated with 150 µl of 1 µg/ml peptide (either AP1 or AP2) in PBS overnight at 4°C. Wells were blocked with 200 µl blocking buffer [0.5% (w/v) BSA in PBS] for 1 h at 37°C and washed three times with 0.1% (v/v) Tween 20 in PBS. IgY was serially diluted in blocking buffer from 100 µg/ml to 0.78 µg/ml. Primary antibody (100 µl) was incubated for 2 h at 37°C. The plate was washed as before prior to addition of 120 µl of rabbit anti-IgY-HRPO conjugate (1:12 000 in blocking buffer) and incubated for 1 h at 37°C. After washing as described, ABTS/H_2O_2 substrate solution [0.05 %(w/v) ABTS, 0.0015% (v/v) H_2O_2 in 150 mM Citrate Phosphate Buffer, pH 5.0] was added and the plate incubated in the dark for 10 min prior to reading the absorbance at 405 nm.

Peptide affinity chromatography columns were prepared for further purification of pools of isolated chicken and-AP1 and AP2 peptide IgY. The AP1 peptide was coupled to an Aminolink® matrix (Pierce, Rockford, IL, USA) according to the manufacturer's instructions. Briefly, the gel slurry (3 ml) was drained and washed with 3 ml of IgY buffer. Peptide (4 mg) was dissolved in IgY buffer, 100 µl of 1 M $NaCHBH_3$ added, and mixed in an end-over rotator for 2 h at RT. The gel slurry was drained, washed with 2 ml of 1 M Tris-Cl (pH 7.4) resuspended in Tris-Cl buffer, pH 7.4 and 100 µl of 1 M $NaCHBH_3$ added before incubation for 30 min at RT. The resin was washed with 10 ml of 1 M NaCl, 10 ml 0.05% (w/v) NaN_3, and finally stored in IgY buffer at 4°C until use.

The AP2 peptide was coupled to a SulfoLink® matrix. The gel slurry (2 ml) was washed with 6 ml of 50 mM Tris-Cl, pH 8.5, 5 mM EDTA. Peptide was added to the slurry and mixed in an end-over rotator for 15 min at RT which was subsequently allowed to settle for 30 min at RT. The column was washed with 3 ml of 50 mM Tris-Cl, pH 8.5, 5 mM EDTA prior to incubation with 50 mM cysteine (1 ml) to block unreacted sites (end-over-end rotator, 15 min, RT). This was left to stand (30 min, room temperature). The column was washed with 56 ml of 1 M NaCl, followed by 2 ml of 0.1 M NaH_2PO_4, 0.05% (w/v) NaN_3, pH 6.5. The resin was stored at 4˚C until required.

Polyclonal anti-AP1 and AP2 peptide IgY was affinity purified using the respective peptide affinity resins prepared as described. Briefly, the column was washed with IgY buffer and the pooled, filtered IgY sample passed over the column once. The IgY sample was subsequently cycled through the column overnight in a reverse direction at 4℃ to avoid blocking of the column and achieve maximum binding of specific antibodies. The column was washed with IgY buffer and eluted with 100 mM glycine-Cl buffer (pH 2.5). Fractions (1 ml) were collected into tubes containing 100 µl 1 M Na-phosphate buffer, pH 8.5 and mixed gently to neutralise the pH. Absorbance of the fractions was measured at 280 nm and fractions containing antibody were pooled. Affinity purified IgY was tested in an ELISA as described above.

5.2.4 Immunisation and challenge of BALB/c mice

Three groups of four mice each were used for an immunisation and infection experiment. The first group was immunised subcutaneously with 10 µg of purified recombinant refolded AP1 with Freund's complete adjuvant for the first immunisation and with Freund's incomplete adjuvant for the three subsequent boosters, two weeks apart. The second group was similarly immunised with 10 µg purified recombinant refolded AP2, while the third group was immunised with a control (unrelated) antigen (VP4). Prior to immunisation, mice were bled from the tail to obtain pre-immune serum samples. Following immunisation, mice were bled weekly into heparinised capillary tubes, the tubes centrifuged (Haematocrit centrifuge) and packed cell volume (PCV) measured. Following this, the sera were stored at -20℃. After the final immunisation, the mice were allowed to rest for 2 weeks and then infected intra-peritoneally with 1×10^4 trypanosomes (*T. congolense* strain IL1180). Tail blood was monitored for trypanosomes by microscopy every two days prior to the onset of parasitaemia and each day afterwards. PCV was monitored weekly. Pooled sera were used in ELISAs to measure antibody titres, and eventually for immunofluorescence microscopy and western blotting.

5.2.5 Trypanosome culture and RNAi of AP1 and AP2

Trypanosome *in vitro* culture, RNAi transfection and the *in vitro* life cycle was conducted as described by Coustou *et al*. (2010). Briefly, procyclic forms of *T. congolense* (strain IL3000) were cultured in complete MEM [MEM Base powder, 25 mM HEPES (acid form), 26 mM $NaHCO_3$, 0.1 mM hypoxanthine, 20% (v/v) foetal calf serum, 4 mM haemin, 2 mM glutamine, 10 mM proline]. Procyclics could be transformed into epimastigotes by centrifuging (2 000 *g*, 10 min, RT) a procyclic culture (of at least 10^7 parasites/ml), resuspending the cells in complete MEM without foetal calf serum, and growing them at 27°C for 1-2 h. The media were subsequently supplemented with 10% (v/v) foetal calf serum and half the medium removed every two days to maintain the epimastigotes in culture. After at least four days metacyclics started appearing, and the mixture of epimastigotes and metacyclics were injected into NOD (NOD.Cg-Prkdescid Il2rgtm1 Wjl/Szj) mice and allowed to proliferate for one to two weeks while tail blood was monitored daily to determine the level of parasitaemia. When parasitaemia reached a minimum of 10^8 parasites/ml, 5 drops of tail blood were collected into 1 ml of BSF adaptation medium [MEM Base powder, 25 mM HEPES (acid form), 26 mM $NaHCO_3$, 5.5 mM D-glucose, 1 mM sodium pyruvate, 40 µM adenine, 0.1 mM hypoxanthine, 20 µM thymidine, 20 µM bathocuproin, 20% (v/v) fresh goat serum, 2 mM glutamine, 5% (v/v) serum plus]. Most red blood cells were removed by centrifugation (200 *g*, 1 min, RT) and the supernatant with parasites was incubated in 24 well plates at 34°C in a humidified atmosphere containing 5% CO_2.

For RNAi of AP1 and AP2 of *T. congolense*, three constructs were created using the $p2T7^{Ti}$ vector (Figure 5.3). Firstly, a 300 bp section of AP1 was amplified from *T. congolense* genomic DNA using primers incorporating an XbaI restriction enzyme site (underlined), AP1XbaFw: 5' AAA TCT AGA AAG GTG CAC CGA GCC GGG TTG C 3' and AP1XbaRv: 5' AAA TCT AGA GTG GTG AGT GGC GTT GTA 3'. A similar set of primers following the same model was used to amplify a 400 bp segment of AP2, AP2XbaFw: 5' AAA TCT AGA CAA TGC AGT GGT GGA AAG AGC 3'; AP2XbaRv: 5' AAA TCT AGA GTG GTG AGT GGC GTT GTA GC 3'. Additionally, a third construct (AP1+AP2) was created using primers allowing the concatenation of the two segments of AP1 and AP2: Ap1Rv: 5' GCT CTT TCC ACC ACT GCA TTG GTG GTG AGT GGC GTT GTA GC 3' used AP1XbaFw and AP2Fw: 5' GCT ACA ACG CCA CTC ACC ACC AAT GCA GTG GTG GAA AGA GC 5' used AP2XbaRv (Figure 5.2). The product of each of these two reactions was purified, as described (see Section 5.2.1), and used together as the template for a PCR with primers AP1XbaFw and AP2XbaRv, resulting in a 700 bp

product. PCR, cloning and transformation into *E. coli* BL21 cells for each of these three amplicons (AP1, AP2 and AP1+AP2) was performed as described (Section 5.2.1). The p2T7^{Ti} vector (LaCount *et al.*, 2000), was prepared for RNAi by restriction digest using XbaI followed by dephosphorylation using SAP. The vector contains an ampicillin resistance gene facilitating selection in bacteria. All constructs were verified by sequencing prior to transformation.

For transformation of procyclic trypanosomes, the method described by Coustou *et al.* (2010) utilising the Amaxa Nucleofaction® transformation system (Lonza) was followed. Briefly, a culture of procyclic *T. congolense* (strain Treu 13:29) parasites (Wirtz *et al.*, 1999) was inoculated in complete MEM media and grown until there were sufficient cells for 1-2 x 10^7 cells/transformation. The strain used, Treu 13:29, had previously been stably transformed with the pLew13 and pLew29 plasmids (Figure 5.3) required for tetracycline regulation of RNAi (Inoue *et al.*, 2002). Prior to transformation, the recombinant p2T7^{Ti} plasmid was isolated by miniprep (as described in Section 5.2.1) and linearised using NotI.

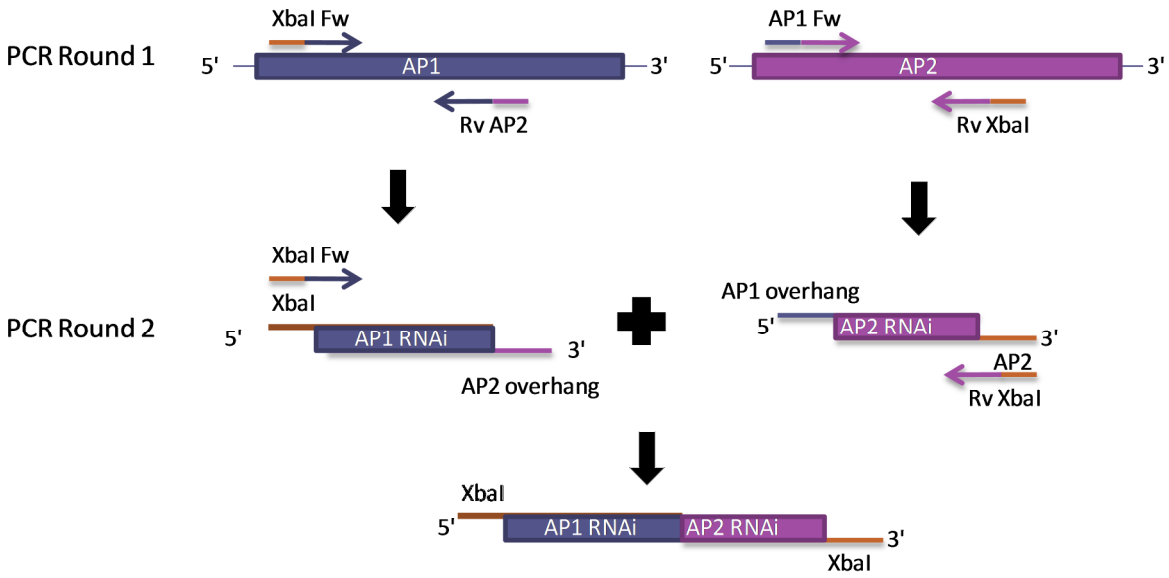


Figure 5.2 Schematic diagram of the two rounds of PCR necessary for the production of a concatenated AP1+AP2 PCR product with added XbaI overhangs for cloning into p2T7^{Ti} for RNAi. For the first round of PCR, a forward primer with a 5' XbaI site and reverse primer with a 3' AP2 overhang was used to amplify a section of AP1 using *T. congolense* genomic DNA as a template. Similarly, a forward primer with a 5' AP1 overhang and a reverse primer were used to amplify a section of AP2. The products of these two reactions were subsequently used as the template for a second round of PCR using the AP1 forward primer and AP2 reverse primer, generating a concatenated product with XbaI overhangs added from the primers.

The digested DNA was checked on a 1% (w/v) agarose gel for complete linearisation and subsequently precipitated using 100% (v/v) ethanol (2.5 volumes) and 3 M sodium acetate

buffer, pH 5.2 (0.1 volumes). Following incubation at -70°C for 2 h, the samples were centrifuged (10 000 *g*, 30 min, 4°C) and washed with 70% (v/v) ethanol. The resulting pellets were air-dried for at least 20 min at 37°C under sterile conditions and resuspended in sterile water to a final DNA concentration of 500 ng/µl. DNA concentration was estimated using OD_{280}.

The required number of parasites was centrifuged (2 000 *g*, 10 min, RT) and resuspended in Amaxa Basic Parasite Nucleofactor® 2 buffer (100 µl/ transfection). Linearised DNA (10 µg) was added and the sample transferred to an electroporation cuvette avoiding bubbles and with minimum pipetting. The sample was subsequently electroporated (program: X-001) and transferred to a 15 ml tube containing 11 ml of pre-warmed complete MEM. The tube was incubated at 27°C for 30 min prior to transferring into a 25 cm^2 flask for incubation overnight at 27°C. The culture was examined under a microscope for the presence of dividing cells, and subsequently divided into 6 wells of a 24 well plate and serially diluted across a 96 well plate with medium containing bleomycin (2 µg/ml). All subsequent culturing was conducted in the presence of bleomycin. Transformation of the parasites with the $p2T7^{Ti}$ recombinant plasmid was confirmed by PCR amplification of the bleomycin gene after isolation of genomic DNA from each transfected strain as well as the wild type strain.

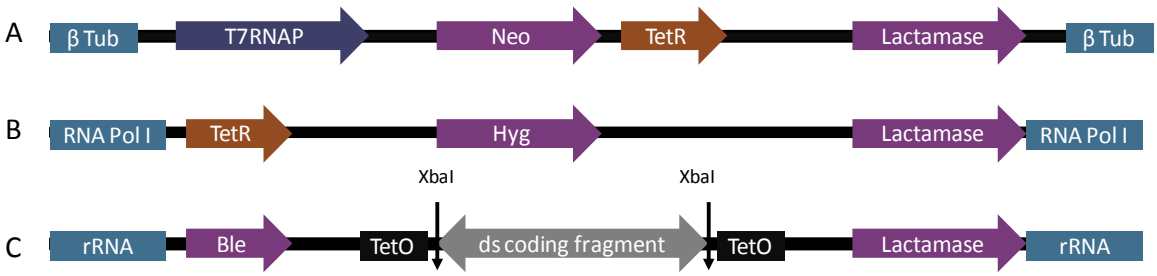


Figure 5.3 Schematic diagram showing pertinent regions of the three vectors necessary for tetracycline induced RNAi in *Trypanosoma congolense*. (A) pLew 13 (B) pLew 29 (C) $p2T7^{Ti}$. Arrows indicate direction of transcription. Blue boxes show regions homologous to genomic DNA as points of insertion for the plasmids (A) β tubulin, (B) RNA Polymerase I, (C) rRNA spacer region. Brown arrows show genes needed for the tetracycline repressor protein (TetR) which prevents transcription of the RNAi fragment shown in (C) in the absence of tetracycline. Dark blue arrow shows the T7 RNA polymerase (T7RNAP) necessary for transcription. Purple arrows indicate genes whose products confer antibiotic resistance to (A) neomycin (Neo), (B) hygromycin (Hyg), (C) phleomycin (Ble), (A,B,C) ampicillin (lactamase). Grey double headed arrow indicates position in which the gene fragment of interest is cloned (using two XbaI restriction sites). (C) Black box (TetO) indicates the tetracycline operator regions to which the TetR protein binds in the absence of tetracycline inhibiting transcription. Diagram adapted from Wirtz *et al.* (1999) and Inoue *et al.* (2002)

Repression of the targeted gene(s) was analysed using western blotting, immunofluorescence and activity assays. For the western blotting, 5 x 10^5 parasites were collected daily for 8 days from cultures of tetracycline (1 µg/ml) induced (Figure 5.4) transfected strains (AP1, AP2 and AP1+2) cultures and the wild type strain (Treu 13:29). Parasites were washed in PBS, counted using a haemocytometer and lysed in 2% (w/v) boiling SDS with the addition of a peptidase inhibitor cocktail (Roche). Reducing SDS-PAGE gels [10% (w/v)] were prepared and 5 x 10^5 parasites loaded/lane. Blots were stripped (Restore Western Blot stripping buffer) and re-probed with either chicken anti-BiP69 IgY (1/30 000) or mouse anti-tubulin IgG (1/2 000) as loading controls to verify that equal quantities of sample were present in each well. ECL immunodetection reagent (GE Healthcare) was used as a substrate. The activity assays were conducted as described in Section 5.2.2, except for the addition of 1 µg of parasite lysate/500 µl instead of purified protein. Activity against 20 µM H-Ala-AMC was measured. Addition of puromycin (1 mM) abrogated activity (data not shown), confirming that catalysis of H-Ala-AMC was due to the alanyl APs and not any other enzymes.

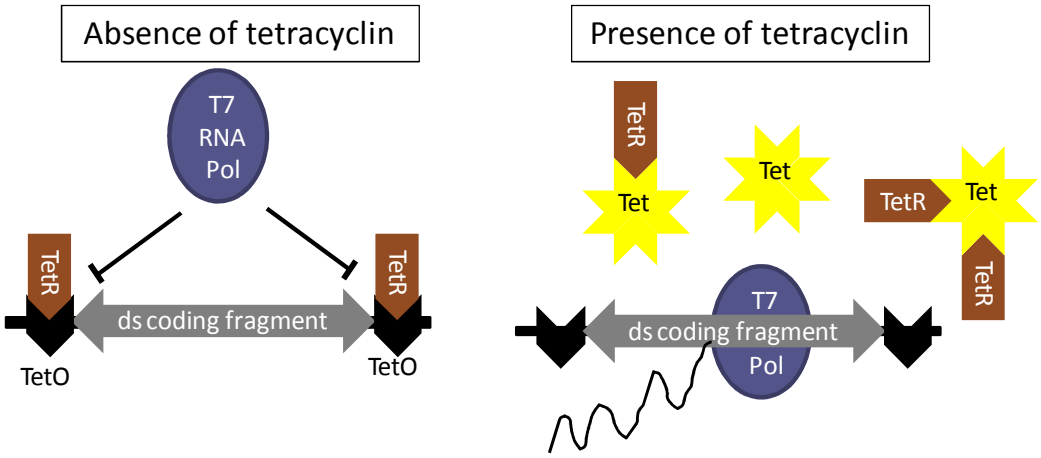


Figure 5.4 Schematic diagram showing the mechanism of the tetracycline regulation of double stranded RNA for RNAi. TetR (tetracycline repressor protein) is encoded by both the pLew13 and pLew29 vectors. TetO (tetracycline operator) found on either side of the multiple cloning site of the p2T7^{Ti} vector binds the TetR in the absence of tetracycline inhibiting transcription by the T7 RNA polymerase encoded by the pLew13 vector. Upon addition of Tet (tetracycline), the TetR binds to the Tet allowing transcription of the fragment cloned into p2T7^{Ti}.

5.2.6 Immunofluorescence microscopy of trypanosomes

Bloodstream form trypanosomes were isolated from mouse blood as described (Lanham and Godfrey, 1970) in Section 2.2.1. For cultured trypanosomes (procyclic, metacyclic and BSF), the parasites were collected by centrifugation (2 000 *g*, 10 min, RT) and washed in 500 µl PSG (Lanham and Godfrey, 1970). The parasite pellet was resuspended in 350 µl 1% (v/v) formaldehyde in PBS and incubated at RT for 10 min,

followed by the addition of 40 µl of 1 M glycine in PBS and further incubation at RT for 10 min. The parasites were then diluted to an appropriate concentration (by checking a drop on a slide) using PBS. Parasites (20 µl/slide well) were added to the slide and allowed to dry for 1 h.

Each slide well was washed twice over 5 min [50 µl PBS with 0.5% (w/v) BSA]. Primary antibody (20 µl/ slide well) was added at an appropriate dilution in PBS containing 0.5% (w/v) BSA and 1% (v/v) Triton X-100, and incubated for 1 h at RT in a humidified atmosphere to prevent evaporation. For affinity purified chicken anti-AP1 IgY, chicken anti-AP2 IgY and chicken anti-C2 IgY, dilutions of 1/20 000 were used. For polyclonal mouse anti-AP1 serum and mouse anti-AP2 serum, dilutions of 1/1000 were used. The dilutions used here were established after testing a range of different concentrations of each antibody, from 1/200 to 1/30 000. Following incubation in primary antibody, the slide wells were washed [50 µl PBS with 0.5% (w/v) BSA] three times over 15 min.

Secondary antibody (20 µl/ slide well) was diluted at 1/100 in PBS containing 0.5% (w/v) BSA and 1% (v/v) Triton X-100, and incubated for 30 min. Secondary antibodies used included Alexa Fluor goat anti-mouse IgG 488, Alexa Fluor goat anti-chicken IgG 488 and Texas Red® goat anti-mouse IgG (Invitrogen). Following a washing step [50 µl PBS with 0.5% (w/v) BSA three times over 15 min], DAPI (20 µl/ slide well) diluted in PBS (final concentration of 0.5 µg/ml) was added to each slide well and incubated at RT for 10 min. Following a washing step, 1 drop of *SlowFade*® Gold anti-fade reagent (Invitrogen) was added to each slide well prior to covering with a coverslip and sealing with nail varnish. Slides were viewed using a Zeiss Axio Imager Z1 fluorescent microscope and images captured using the MetaMorph® software at a total magnification of 100x.

5.3 RESULTS

5.3.1 Cloning and expression of AP1 and AP2

Genomic DNA was successfully isolated from *T. congolense* cells at a concentration of 769 ng/µl as determined by spectrophotometery, in a final volume of 100 µl. As can be seen from Figure 5.5 A, there is a single band of indeterminate size with a slower electrophoretic mobility than the 10 kb band. No contaminating RNA is visible on the gel. Figure 5.5 B shows the PCR products of the expected size of 2.5 kb produced upon amplification of the AP1 and AP2 genes using the genomic DNA as a template.

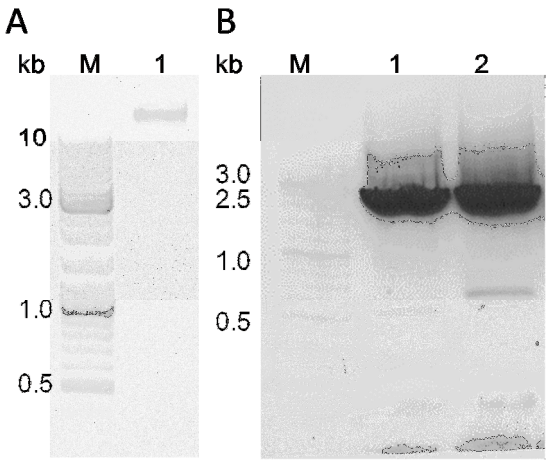


Figure 5.5 Agarose gel analysis of *T. congolense* genomic DNA isolation and PCR amplification of *AP1* and *AP2* genes. (A) Genomic DNA isolation from *T. congolense.* Lane 1: purified genomic DNA (2 µl). (B) AP1 and AP2 PCR products. Lane 1: PCR product of AP1; lane 2: PCR product of AP2. M: GeneRuler® kb DNA ladder (1 µl).

Expression of AP1 and AP2 using *P. pastoris* vectors pPic9 (in strain GS115) and pPicZα28 (in strain X-33) was not successful, regardless of changes in media (BMM and BMMY), buffering pH (6.0 - 7.5), quantity of methanol [0.1 – 1% (v/v)] used for induction and expression time (2 – 10 days). Expression of AP1 and AP2 in pGEX-4T-1, pET28a and pET32a all gave similar expression patterns in *E. coli* (BL21 DE3) as insoluble inclusion bodies and expression was henceforth continued using only the His-tagged pET32a vector. Figure 5.6 shows the 91 kDa bands of AP1 and AP2 expressed upon induction with 0.4 mM IPTG. The expressed protein was found only in the insoluble fraction of the cell lysate.

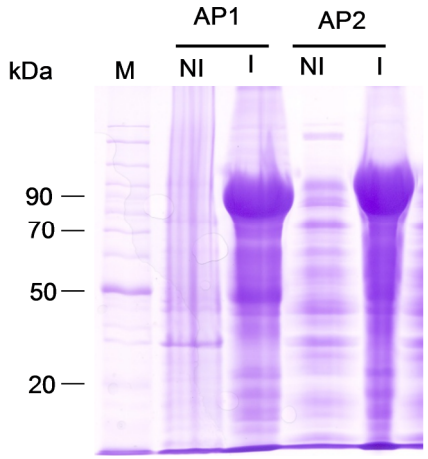


Figure 5.6 Reducing 10% SDS-PAGE showing expression of AP1 and AP2 in *E. coli* (BL21 DE3). M: BenchMark® protein ladder (Invitrogen); NI: non-induced *E. coli* culture; I: expression induced by addition of 0.4 mM IPTG.

To evaluate if it were possible to solubilise the expressed aminopeptidases, increasing amounts of urea was added to a sample of cell lysate after AP1 expression. As evident

from Figure 5.7, very little of the protein of interest was solubilised even using high concentrations of urea since very little protein was found in the supernatant fraction (Figure 5.7 B), while the majority of the protein was still in the insoluble fraction (Figure 5.7 A). The greatest amount of AP1 was solubilised with 2 M urea (Figure 5.7 B) and further addition of urea did not increase the solubilisation. The solubility of AP2 in urea followed a similar pattern.

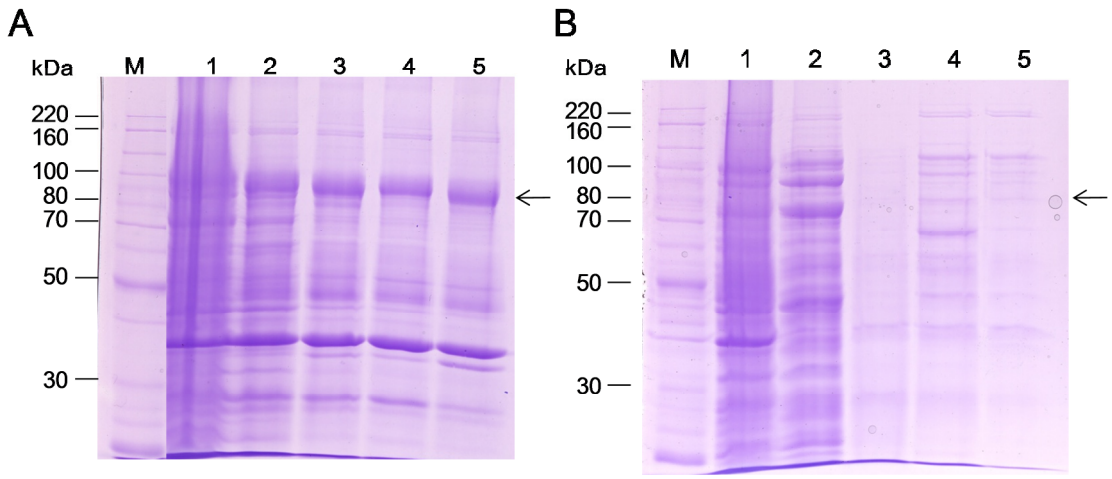


Figure 5.7 Reducing 10% SDS-PAGE analysis of solubilisation of AP1 with urea. (A) Pellets from induced expression of AP1 in *E. coli* (BL21DE3) Lane 1: Non-induced; lane 2: 2 M urea; lane 3: 4 M urea; lane 4: 6 M urea; lane 5: 8 M urea. (B) supernatant from induced expression of AP1 in *E. coli* (BL21DE3) using the same concentrations of urea as (A). Lane 1. M: BenchMark® protein ladder (Invitrogen). Arrows indicate protein of interest.

Refolding of AP1 and AP2 was attempted using a variety of buffers with different concentrations of guanidine, L-arginine and different ratios of oxidised to reduced glutathione. Table 5.1 shows the results obtained from the nine different buffers used with AP1 and AP2 lysate. The presence of a precipitate indicates that some protein is insoluble. However, as the refolding was done on the incompletely purified inclusion bodies and not purified protein, it is possible that the precipitated protein observed was in fact not the protein of interest. This seems to be the case with Buffer 1 for AP2 where, even though a significant amount of precipitate was observed, the highest level of activity against H-Ala-AMC was found. Following this trial, Buffer 2 was routinely used for refolding of AP1, while Buffer 1 was used for AP2.

Table 5.1 Refolding buffers with variable concentrations of guanidine, L-arginine, and different reduced glutathione to oxidised glutathione ratios gave different levels of refolding of AP1. Relative activity was measured by the hydrolysis of H-Ala-AMC. Base refolding buffer: 55 mM Tris-Cl buffer, 21 mM NaCl, 0.88 mM KCl, pH 8.2.

Buffer	[Guanidine] (M)	[L-Arg] (M)	[GSH:GSSG]	Presence of precipitate		Relative activity against H-Ala-AMC (%)	
				AP1	AP2	AP1	AP2
1	0.4	0	2:0.2	+++	+++	0	100
2	0.4	0.4	2:0.4	+	+	100	51
3	0.4	0.8	1:0.1	-	+	0	0
4	0.4	0	2:0.4	-	++	0	27
5	0.9	0.4	1:1	-	+	11	39
6	0.9	0.8	2:0.2	+++	-	0	0
7	1.5	0	1:1	-	+	78	0
8	1.5	0.4	2:0.2	-	-	8	0
9	1.5	0.8	2:0.4	+	-	90	0

+++ large ppt; + minimal precipitate; - no precipitate

In order to verify that the AP proteins were indeed correctly refolded, the products from four different refolding buffers were analysed by reducing and non-reducing SDS-PAGE and representative results obtained for AP1 are showed in Figure 5.8. The groups of reduced and non-reduced AP1 samples in Figure 5.8 clearly shows a change in size between the two. Those lanes containing reduced samples (lanes 1-4) migrated corresponding to a larger molecular weight than those which were not reduced (lanes 5-8). However, the low concentration of the protein of interest and the presence of contaminating bands, make it difficult to identify differences between refolding buffers.

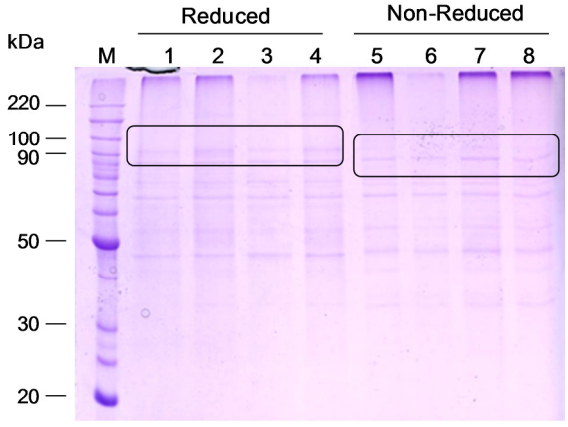


Figure 5.8 10% SDS-PAGE showing the change in size of renaturated AP1 before and after reducing with β-mercaptoethanol. M: BenchMark® protein ladder (Invitrogen). Lanes 1-4: samples reduced by 0.7% mercaptoethanol and boiled; lanes 5-8: samples not reduced, nor boiled. Lanes 1 and 5: refolding buffer 1; lanes 2 and 6: refolding buffer 2; lanes 3 and 7: refolding buffer 3; lanes 4 and 8: refolding buffer 4. See Table 5.1 for details of refolding buffers.

After refolding is was necessary to further purify the recombinant APs and this was achieved using a His-affinity column (Figure 5.9). As can be seen in Figure 5.9 B, the peak of OD_{280} in Figure 5.9 A corresponds to the elution of AP2. It is also evident that the protein is not entirely pure as there were several low molecular weight contaminating proteins.

For this reason it was necessary to further purify the eluate using a Q-Sepharose column. However, due to the high concentration of NaCl used in the His-affinity column elution buffer (0.5 M), the fractions containing the protein of interest were pooled and diluted 3 fold in preparation for the ion-exchange column. A similar profile was obtained for the His-affinity purification of AP1.

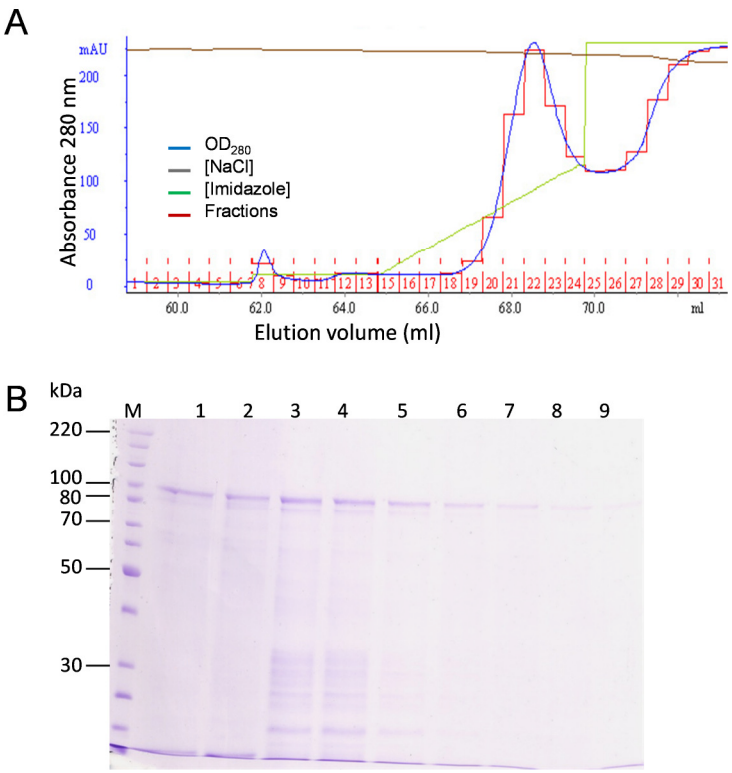


Figure 5.9 Partial purification of AP2 using a His-affinity resin. (A) Elution profile: linear gradient of 50 to 500 mM imidazole in 5 ml followed by step increase to 1M imidazole in 10 ml. (B) 10% Reducing SDS-PAGE of fractions 19-27 (lanes 1-9) eluted from the His-affinity column. M: BenchMark® protein ladder (Invitrogen).

Figure 5.10 shows the elution profile from a Q-Sepharose ion exchange column. Only a single band at 91 kDa corresponding to AP2 is present in some fractions (Figure 5.10 B), indicating high levels of purification, albeit at very low concentration of protein. Also, no bands are visible in lane 2 of Figure 5.10 B, indicating that all the protein present in the

sample bound to the column. AP2 is eluted in a single narrow peak at approximately 300-400 mM NaCl (Figure 5.10 A).

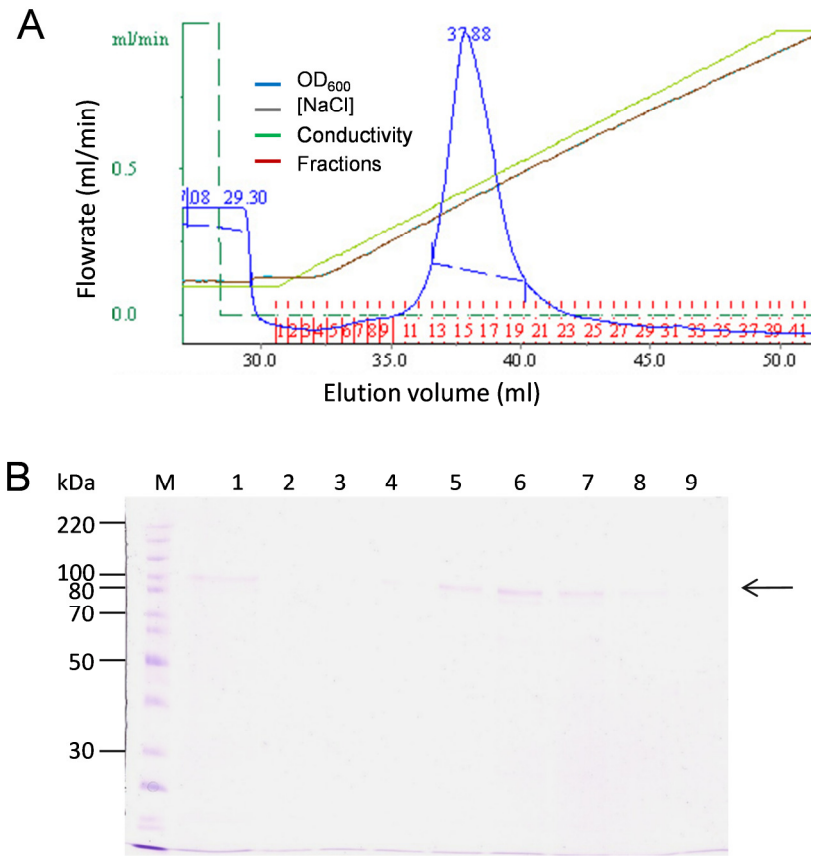


Figure 5.10 Complete purification of AP2 using Q-Sepharose ion exchange chromatography. (A) Elution profile: linear gradient of 160 mM to 1 M NaCl in 20 ml. (B) 10% Reducing SDS-PAGE. Lane 1: Pooled and dialysed fractions from His elution; lane 2: non-retained fraction; lanes 3 to 9: fractions 12 -18 eluted with 300 to 400 mM NaCl. M: BenchMark® protein ladder (Invitrogen). Arrow indicates the band of interest.

Prior to immunisation, it was necessary to remove the His-tag from the recombinant refolded AP1 and AP2. Due to the low stability of the APs at RT, it was necessary to determine the optimum length of time required for thrombin cleavage of the His-Tag with minimal degradation of the APs. Thrombin was added to an aliquot of AP1 and incubated at RT for 2, 4 and 6 h (Figure 5.11 lanes 1-3). As a control, an equal quantity of AP1 was incubated at RT without thrombin to determine the rate of protein degradation (Figure 5.11, lanes 4-6). It was subsequently determined that at least 6 h at RT was necessary for complete cleavage of the His-tag, even though some degradation of AP1 did occur in 6 h. A similar result was obtained for AP2.

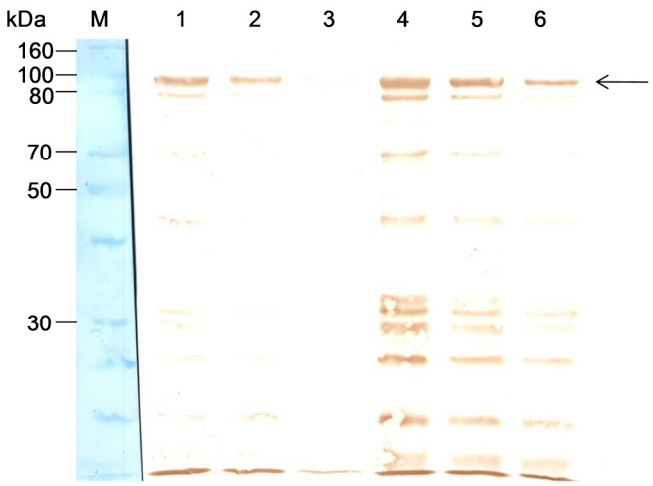


Figure 5.11 Western blot to determine optimum period of His-tag cleavage by thrombin with minimal degradation of AP. 10% reducing SDS-PAGE gel was blotted at 24V for 1 h using a semi-dry blotter. Lanes 1-3: AP2 and 1 unit thrombin; lanes 4-6: AP2 alone. Lanes 1 and 4: incubated for 2 h at RT; lanes 2 and 5: incubated for 4 h; lanes 3 and 6: incubated for 6 h. M: BenchMark® protein ladder (Invitrogen). Arrow indicates band of interest.

5.3.2 Enzymatic characterisation

The substrate specificities of the recombinant, refolded APs were investigated (Table 5.2). It is evident that the APs have a well defined specificity for amino acids with small, non-polar R-groups especially given the fact that no catalysis was observed for the substrates H-Arg-AMC or H-Tyr-AMC. Furthermore, it seems that the APs under investigation are indeed Family M1 type alanyl aminopeptidase since the highest specificity constant (by more than 4 fold) was established for the substrate H-Ala-AMC. In comparison to each other, AP1 and AP2 have very similar specificities for the different substrates tested, with no significant differences found between the two enzymes.

Table 5.2 Kinetic constants for the hydrolysis of synthetic peptide substrates by recombinant refolded *T. congolense* aminopeptidases AP1 and AP. Values calculated using the means of triplicate experiments. K_m was calculated using the Hyper32 software. Assays were carried out in assay buffer [50 mM Tris-Cl buffer, pH 8.0; 1 mM $MnCl_2$] with 15 ng of enzyme at 37°C. No hydrolysis of H-Arg-AMC and H-Tyr-AMC was observed after 4 h incubation at 37°C.

Substrate	Enzyme	V_{max} (s)	K_m (µM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}.mM^{-1}$)
H-Leu-AMC	AP1	0.1092	32.9	0.07	2.21
	AP2	0.0525	48.62	0.04	0.83
H-Gly-AMC	AP1	58.1	82.95	38.73	466.95
	AP2	51.71	57.9	39.78	686.99
H-Ala-AMC	AP1	377.1	139.1	251.40	1807.33
	AP2	236.2	98.02	181.69	1853.62
H-Met-AMC	AP1	6.323	86.10	4.22	48.96
	AP2	14.98	164.80	11.52	69.92

The influence of various divalent cations at two different concentrations on the activity of AP1 and AP2 was analysed (Table 5.3). The similarities between the two concentrations of metal cations suggest that the lower concentration (1 mM) was sufficient for the highest effect on activity on the enzyme, thus addition of ten times more metal cation made no difference to the activity. Furthermore, apart from the instance of Zn^{2+} AP1 and AP2 were affected in a similar manner by all the metal cations tested. Cu^{2+} completely abolished activity of AP2 and decreased the activity of AP1 by more than 50%. Co^{2+}, Fe^{2+}, Ni^{2+}, Mn^{2+} and Mg^{2+} were all similar in increasing the activity of AP1 and AP2 when compared to the absence of a metal ion. Interestingly, the addition of Zn^{2+}, slightly decreased the activity of AP1 but increased the activity of AP2.

Table 5.3 Influence of divalent metal cations on the activity of recombinant refolded AP1 and AP2 of *T. congolense*. The substrate H-Ala-AMC was used to measure activity of 1.5 ng enzyme after a 30 min incubation at 37℃ in assay buffer containing either 1 or 10 mM of the different metal ion salts. Data points reflect the means of triplicate experiments ±SE.

Metal	Concentration (mM)	Relative Activity (%)	
		AP1	AP2
No cation	0	(100)	(100)
Ca (II)	1	385 ± 4	216 ± 8
	10	121 ± 5	137 ± 17
Co (II)	1	157 ± 3	141 ± 1
	10	151 ± 7	84 ± 5
Cu (II)	1	46 ± 4	0
	10	12 ± 8	0
Fe (II)	1	128 ± 4	121 ± 13
	10	139 ± 12	126 ± 4
Ni (II)	1	133 ± 4	146 ± 6
	10	360 ± 6	141 ± 13
Mg (II)	1	108 ± 6	138 ± 7
	10	107 ± 9	149 ± 6
Mn (II)	1	131 ± 5	135 ± 1
	10	140 ± 5	158 ± 7
Zn (II)	1	88 ± 1	114 ± 5
	10	84 ± 1	117 ± 2

The pH optima for AP1 and AP2 were pH 7.5 and 8.0 respectively (Figure 5.12 A). Both enzymes have a defined peak of activity at neutral pH, with activity declining sharply at both more acidic and more basic pH. In spite of this, the two enzymes were stable over a broad pH range, retaining activity after incubation in buffers from pH 4.0 to 10 (Figure 5.12 B).

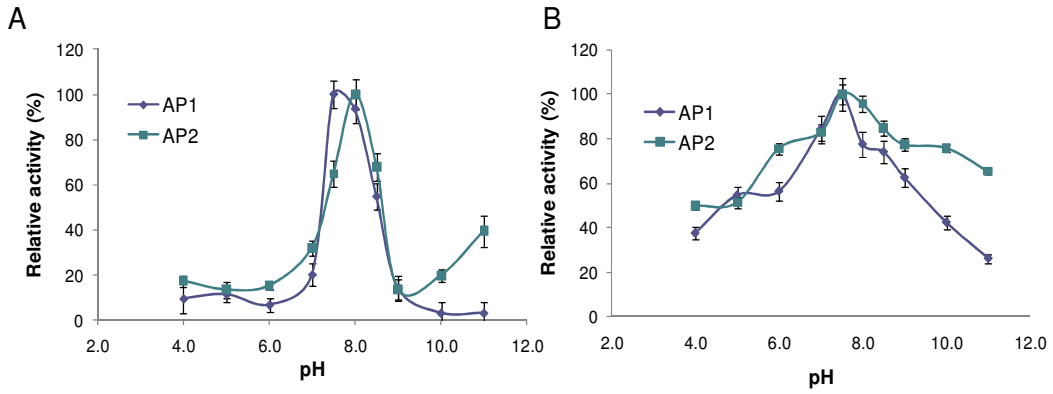


Figure 5.12 pH optimal activity and stability profiles of AP1 and AP2 across the pH range 4.0 to 11.0 using H-Ala-AMC. (A) optimal pH activity profile of enzymes (1.5 ng) assayed for activity after incubation in AMT buffers (100 mM Na-acetate, 200 mM Tris-Cl buffer, 100 mM MES, 6 mM DTT, 4 mM Na_2EDTA) of different pH for 5 min at 37°C. (B) pH stability profile of enzymes (1.5 ng) assayed for activity at pH 8.0 and 7.5 (pH optima for AP1 and AP2 respectively) after incubation in AMT buffers of different pH for 2 h at 37°C. Data points reflect the means of triplicate experiments ±SE and where error bars are not visible data points are contained within the symbol.

AP1 and AP2 are metal dependent aminopeptidases. Hence, the addition of metal chelators such as EDTA and 1,10 - phenanthroline inhibit their activity (Table 5.4). It was also found that the antibiotic puromycin, which is used to distinguish between cytosolic alanyl APs and membrane bound APs (Gros *et al.*, 1985), inhibited AP1 and AP2, confirming that the two enzymes are indeed cytosolic APs. Also, the two APs were found to have similar inhibition constants for all the inhibitors tested. Furthermore, the slow binding inhibition of the AP specific inhibitors bestatin and amastatin was measured. It was found that bestatin and amastatin were potent inhibitors of AP1 and AP2 with similar K_i values of 7.015 and 2.46 (AP1) and 3.4 and 8.6 nM (AP2) respectively. Puromycin proved to be the least effective inhibitor of those tested with a K_i of 116 and 447 µM for AP1 and AP2 respectively.

Table 5.4 Inhibition constants of recombinant refolded aminopeptidases AP1 and AP2 of *T. congolense*. Data points reflect the means of triplicate experiments±SE.

Inhibitors	Concentration	Inhibition constant K_i	
		AP1	AP2
EDTA	0.5 mM	0.14 ± 0.05 nM	0.3 ± 0.06 nM
1,10 phenanthroline	0.1 mM	48 ± 8 µM	34 ± 8 µM
DTT	0.5 mM	78 ± 3 µM	65 ± 9 µM
Puromycin	1 mM	116 ± 7 µM	447 ± 13 µM
Bestatin	5 – 50 nM	7.015 ± 0.1 nM	3.4 ± 1.2 nM
Amastatin	0.05 – 1.25 µM	2.46 ± 0.4 nM	8.6 ± 0.6 nM

*No inhibition was observed with AEBSF (1 mM), E-64 (10 µM), oxidised glutathione (1 mM), leupeptin (10 µM) or pepstatin A (10 µM) for either AP1 or AP2.

5.3.3 Anti-AP peptide antibody preparation

The amino acid sequences of AP1 and AP2 were analysed in order to find peptides with sufficient hydrophilicity, surface probability and flexibility to use as immunogens to raise antibodies reactive with the two proteins (Figure 5.13). For AP1 a 13 amino acid long sequence was chosen, corresponding to residues 630 to 642. The sequence was equally hydrophilic, flexible and surface-located at its N- and C-termini and was coupled via the N-terminal Leu residue to the rabbit albumin carrier using glutaraldehyde because of the relative simplicity of the coupling method.

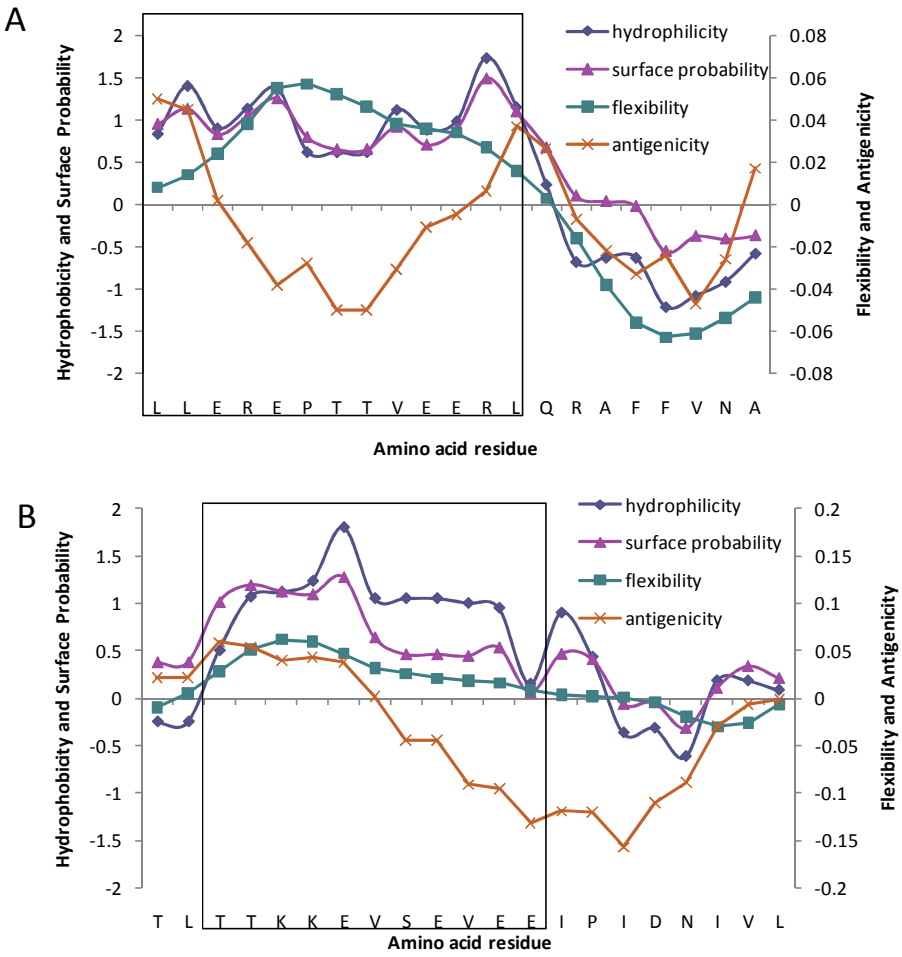


Figure 5.13 Analysis of AP1 (residues 630 to 650) and AP2 (residues 62 to 72) amino acid sequences using Predict7 to select optimal peptides for raising antibodies. (A) AP1 and (B) AP2. Peptides were selected on the basis of four variables: hydrophilicity, surface probability, flexibility, and antigenicity. The boxed regions indicate the peptides use for immunisation.

The algorithm used for prediction of antigenicity described by Welling *et al*. (1985) is based on data collected from 69 continuous epitopes of 20 different proteins. It has been suggested that data from more proteins would need to be studied to improve the prediction of this variable (van Regenmortel and Muller, 1999). Therefore, the low

antigenicity seen in the peptide selected for AP1 (Figure 5.13 A) may not be a reliable prediction. For AP2, an 11 amino acid long sequence, corresponding to residues 62 to 72, was decided to be optimal. The sequence was more hydrophilic and flexible and showed higher values for surface probability towards the N-terminus as compared to the C-terminus, and was consequently coupled to rabbit albumin via a Cys residue added to the C-terminal end.

The peptides selected for anti-peptide antibody production are shown on three dimensional models of AP1 and AP2 in Figure 5.14 A and B respectively. As can be seen from the models, the selected peptides (coloured red in Figure 5.14) of both APs are on the surface of the structure. Also, given the distance of the peptides from the active site of the APs (coloured yellow in Figure 5.14), it is unlikely that the antibodies produced against these peptides would inhibit the activity of the AP.

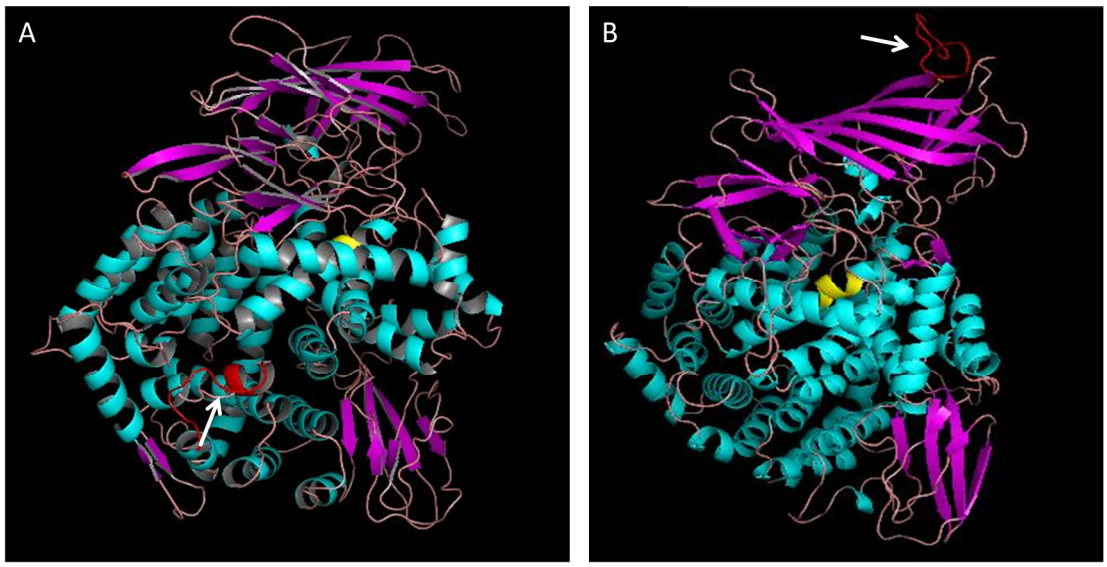


Figure 5.14 Projected structures of (A) AP1 and (B) AP2 showing peptides selected for antibody production. AP sequences modelled on AP N (*E. coli*) using Swiss-Model (Arnold *et al.*, 2006; Kiefer *et al.*, 2009) viewed using Pymol Sequence Viewer (DeLano, 2004). Selected peptide (red, indicated by white arrow) and active sites (yellow) are shown. Secondary structure is indicated by colour: β pleated sheets (purple), α-helices (cyan) and random coils (pink).

After immunisation of chickens with the respective peptide-rabbit albumin conjugates, IgY was isolated from eggs produced by each chicken. In order to determine the titre of the isolated anti-peptide antibodies, ELISAs were performed using the peptide (either AP1 or AP2) to coat the plate. As can be seen from Figure 5.15 all four chickens produced antibodies against the respective peptides. One chicken immunised with the AP1 peptide (Figure 5.15 A) produced a higher titre of antibodies during weeks 8 to 12 than weeks 4 to 7. The second chicken immunised with the AP1 peptide (Figure 5.15 B) produced

antibodies with approximately equally high titres from week 4 until week 12 and the titres were comparable to those of weeks 4-7 of the first chicken. Also important to note is that the pre-immune IgY samples from the chickens immunised with the AP1 peptide (Figure 5.15 A and B) gave a lower background response (absorbance of approximately 0.2) than the samples from the chickens immunised with the AP2 peptide (Figure 5.15 C and D) which showed an absorbance of 0.5-0.8 at 405 nm. For the AP2 peptide (Figure 5.15 C and D) it appears that the antibody titre increases from week 4 to reach a peak at week 7. During the following weeks (8 to 12) antibody titres were not very much higher than that of the pre-immune sample used as a control.

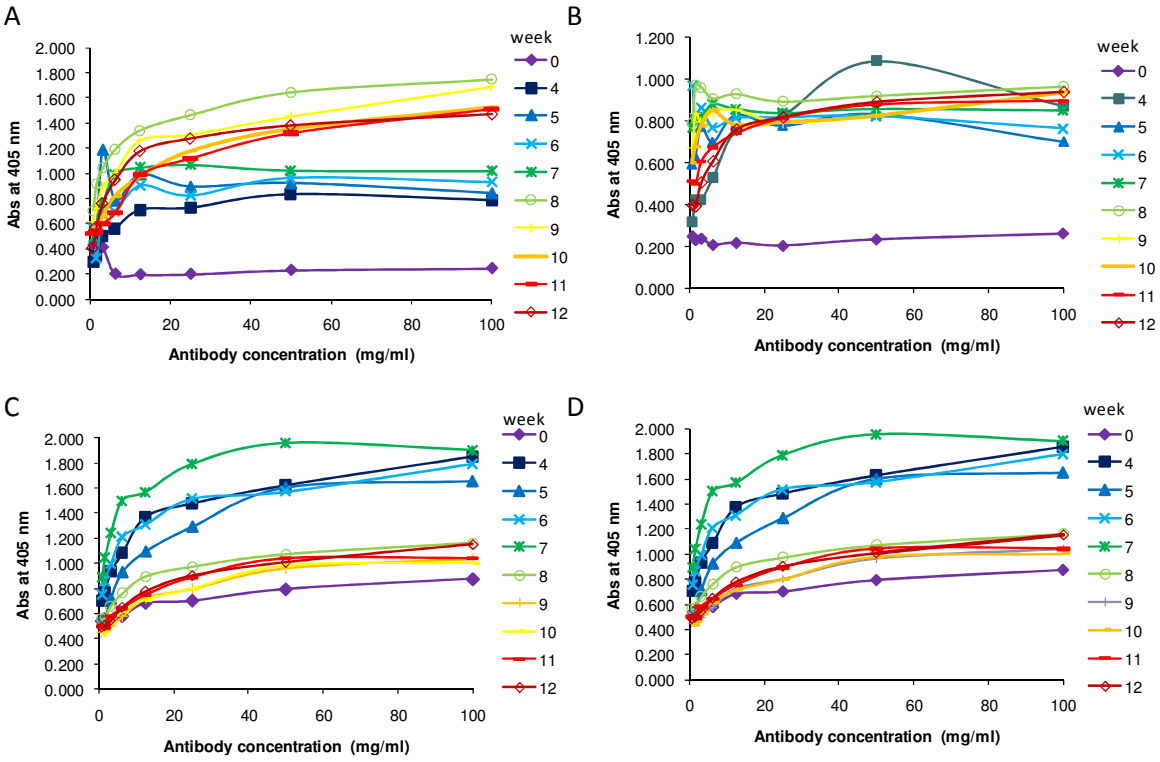


Figure 5.15 ELISAs showing IgY response to immunisation with AP1 and AP2 peptide carrier conjugates. (A, B) ELISA of antibodies produced by chickens 1 and 2 immunised with AP1. (C, D) ELISA of antibodies produced by chickens 3 and 4 immunised with AP2. ELISA plates were coated with peptide (150 µl of 1µg/ml per well). ELISAs were done on IgY pools for each week from two chickens for each peptide. See Section 5.2.3 for details of method.

In order to isolate anti-AP1 or anti-AP2 peptide specific antibodies, it was necessary to perform an affinity purification of the crude IgY. This was achieved by coupling the synthetic peptide used for immunisation for each protein to an affinity resin. Under optimal conditions for binding, only specific anti-AP1 or anti-AP2 peptide antibodies recognise the immobilised peptide while the non-specific antibodies pass straight through the column. The bound anti-peptide antibodies are subsequently eluted using acidic pH and the elution profiles for the various pooled fractions can be seen in Figure 5.16. IgY isolated from eggs collected from weeks 4-8 and weeks 9-12 were separately pooled for

each peptide (i.e. AP1 or AP2) and for each chicken for affinity purification. A comparison of Figure 5.16 A and B with C and D shows that several fractions for anti-AP1 peptide IgY have a much higher absorbance (maximum > 3) and, therefore, a higher concentration of antibodies than anti-AP2 peptide IgY (maximum of 1). Each of the two chickens immunised with the AP1 and AP2 peptides appear to have produced equivalent amounts of antibody. Furthermore, there does not seem to be a significant difference in specific anti-peptide antibody concentration between the pools of IgY produced during weeks 4-8 and weeks 9-12. The affinity purified anti-peptide antibodies were used in subsequent experiments to localise APs in different life cycle stages of *T. congolense* and in RNAi studies.

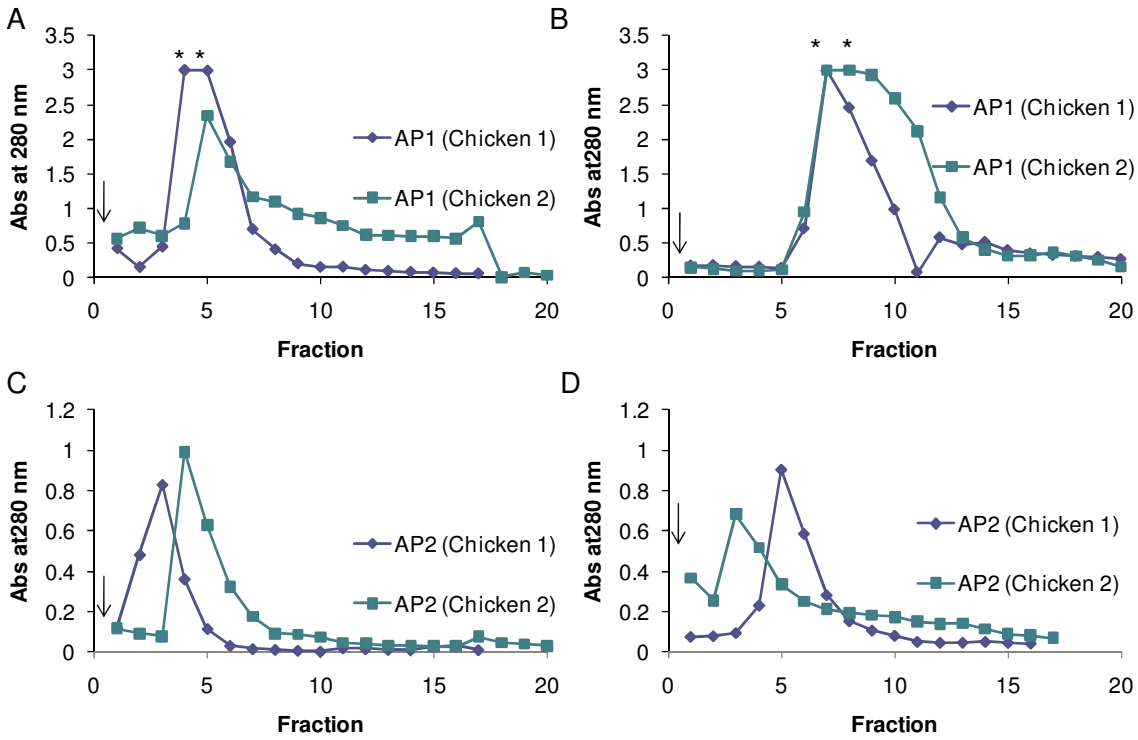


Figure 5.16 Elution profiles of affinity purification of anti-AP1 peptide IgY (A and B) and anti-AP2 peptide IgY (C and D) for weeks 4-8 (A and C) and 9-12 (B and D). See Section 5.2.3 for details of method. Asterisks (*) indicate points measured at the detection limit of instrument. Arrows indicate point at which elution with low pH glycine-Cl buffer was started.

5.3.4 Localisation of native alanyl aminopeptidase

In order to identify whether the alanyl APs are developmentally regulated, lysate of *T. congolense* from three different life cycles stages (procyclics, metacyclic and bloodstream forms) were analysed using anti-AP1 peptide antibodies (Figure 5.17 A) and anti-AP2 peptide antibodies (Figure 5.17 B). A single band at approximately 90 kDa in all lanes in both Figure 5.17 A and B, indicated that the two APs are present throughout the life cycle of *T. congolense* and are not developmentally regulated.

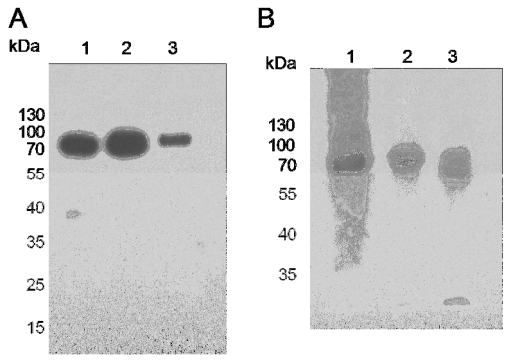


Figure 5.17 Western blot showing presence of AP1 and AP2 in procyclic, metacyclic and bloodstream form *T. congolense* (strain IL3000). Lane 1: Procyclic lysate; lane 2: metacyclic lysate; lane 3: bloodstream form lysate. Primary antibodies were (A) chicken anti-AP1 peptide antibody (1 µg/ml) and (B) chicken anti-AP2 peptide antibody (1 µg/ml). Secondary antibody for both blots was rabbit anti-chicken-HRPO (1/30000). ECL immunodetection reagent was used as a substrate.

A second method was used to verify that the APs are not developmentally regulated, viz. activity assays. Trypanosome lysate from procyclics, epimastigotes, metacyclics and bloodstream forms were tested for activity against H-Ala-AMC (Figure 5.18). In order to show that the cleavage of H-Ala-AMC was indeed catalysed by the APs and not other peptidases, a control was included where puromycin (a specific inhibitor of alanyl APs) was added to the trypanosome lysate prior to the activity assay. As can be seen from Figure 5.18, addition of puromycin almost completely inhibited cleavage of H-Ala-AMC (less than 10% relative activity), indicating that it is indeed the APs responsible for the cleavage of the substrate.

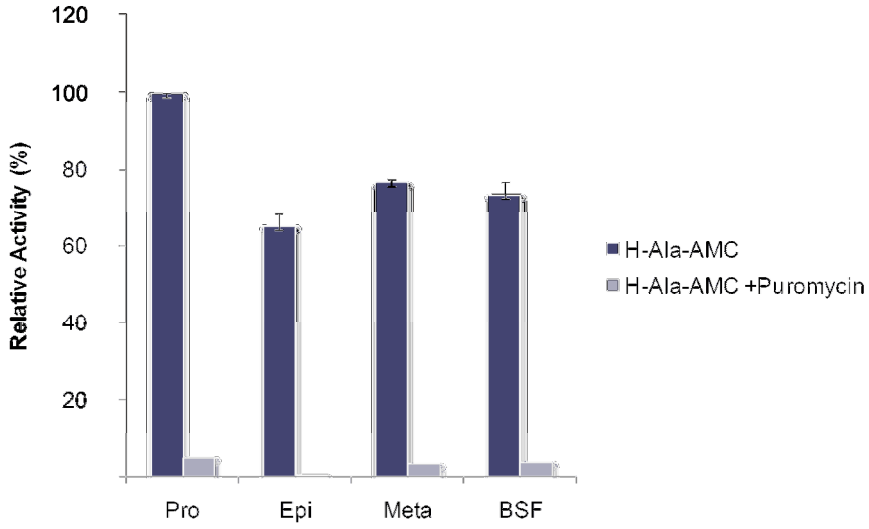


Figure 5.18 Family M1 type aminopeptidase activity in different life-cycle forms of *T. congolense* (strain IL3000). Procyclic (Pro), epimastigote (Epi), metacyclic (Meta) and bloodstream form (BSF) lysates were assayed against H-Ala-AMC. The standard deviations from three experiments are indicated.

In order to establish the initially proposed cytoplasmic location of AP1, immunofluorescence microscopy was performed using mouse anti-AP1 polyclonal antibody (Figure 5.19). Since the APs are not developmentally regulated, it was possible to visualise the APs in procyclic, metacyclic and bloodstream form trypanosomes. AP1 was co-localised with congopain (CP2) that further definitively establishes the life cyclic stage, since congopain is not found in procyclics, is present in small quantities in metacyclics, but is highly expressed in bloodstream forms (Mbawa *et al.*, 1991a), and this can be clearly observed in panels B, G and D respectively. Congopain was detected using an affinity purified chicken anti-peptide antibody, specific for the N-terminal region of congopain and could be observed throughout the cytoplasm of the trypanosome in procyclics, metacyclics and bloodstream forms (Figure 5.19).

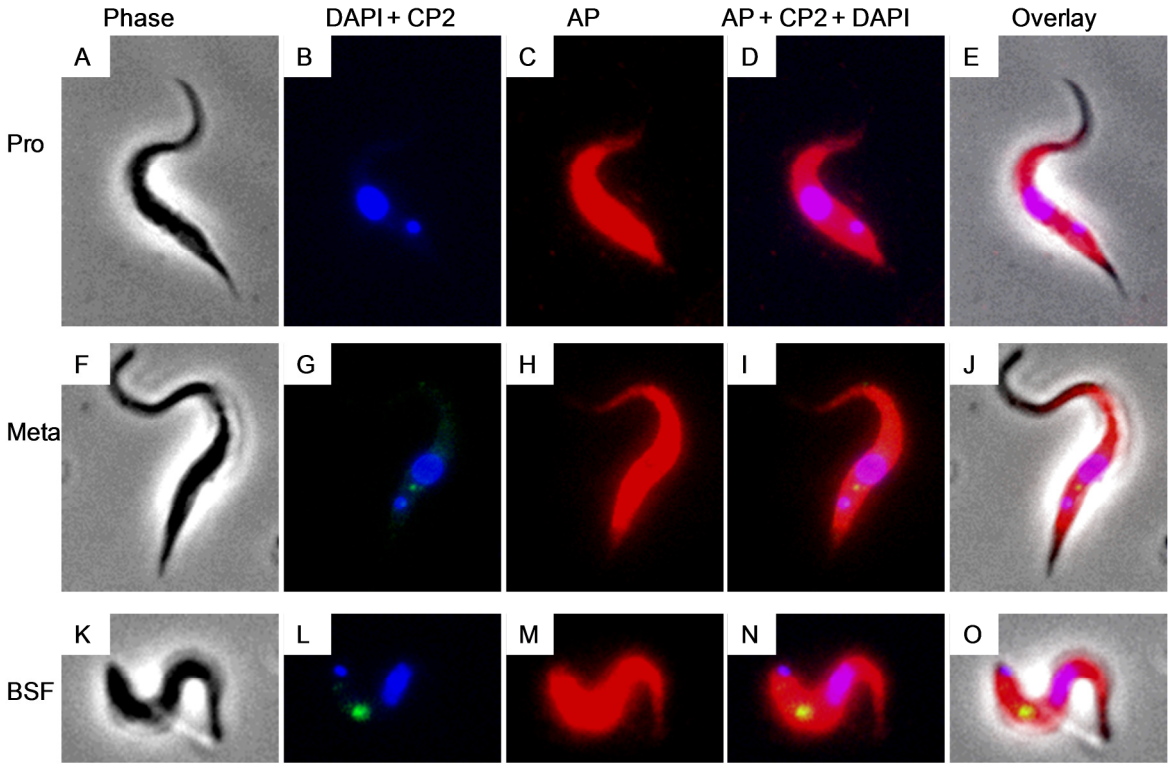


Figure 5.19 Immunofluorescent localisation of AP1 (red) and congopain (green) in *Trypanosoma congolense* procyclics (Pro), metacyclics (Meta) and bloodstream forms (BSF). (A, F, K) Phase contrast of procyclic, metacyclic and BSF trypanosome respectively; (B, G, L) blue (DAPI) and green (congopain) fluorescence overlay of procyclic, metacyclic and BSF trypanosome respectively; (C, H, M) red (AP1) fluorescence of procyclic, metacyclic and BSF trypanosome respectively; (D, I, N) red (AP1), blue (DAPI) and green (congopain) overlay of procyclic, metacyclic and BSF trypanosome respectively; (E, J, O) same as (D) with phase contrast overlay in procyclic, metacyclic and BSF trypanosome respectively. Primary antibodies were mouse anti-AP1 polyclonal (1/1000) and affinity purified chicken anti-peptide (1/20 000) antibodies specific for congopain (CP2). Fluorescent tagged secondary antibodies were Alexa Fluor goat anti-chicken IgG 488 (green) and Texas Red® goat anti mouse IgG (red) at a dilution of 1/1000. See Section 5.2.6 for details of slide preparation.

5.3.5 Immunisation and challenge of BALB/c mice

Three groups of BALB/c mice were used to evaluate the protective potential of *T. congolense* AP1 and AP2. Two groups of mice were immunised with recombinant, refolded protein and the antibody titre measured over 5 weeks by ELISA (Figure 5.20). For AP1 (Figure 5.20 A) it can be seen that the antibody titres were highest at weeks 3 and 5. On the other hand, for AP2 (Figure 5.20 B) the highest antibody titres were obtained at weeks 3 and 4. For both groups of mice, it is evident that the mice responded to the immunisation.

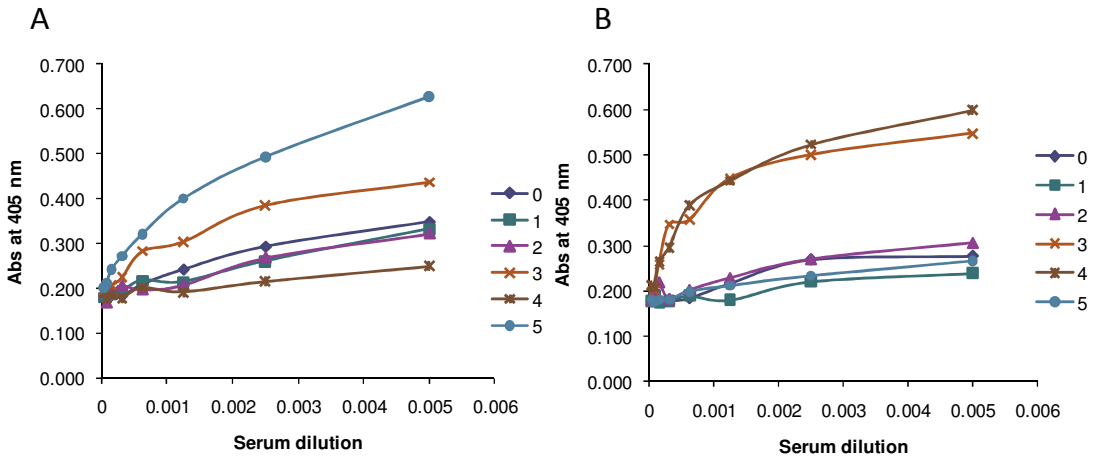


Figure 5.20 ELISA to show immune response of mice after immunisation with recombinant refolded AP1 and AP2. (A) ELISA of pooled serum from mice immunised with AP1. (B) ELISA of pooled serum from mice immunised with AP2. Four mice from each group were bled weekly and the serum pooled and diluted in PBS for the ELISA. The ELISA plates were coated with recombinant refolded AP1 or AP2 (150 µl of 1 µg/ml per well). The secondary antibody was goat anti-mouse IgG-HRPO conjugate (1:12 000). The substrate was ABTS/H_20_2 solution [0.05 %(w/v) ABTS, 0.0015% (v/v) H_2O_2 in PBS].

The immunised mice were rested for two weeks and subsequently challenged by infection with *T. congolense* (IL1180). There was no difference in prepatent period, as mice from all groups had parasites visible in tail blood smears one day after infection. Furthermore, measurement of anaemia by haematocrit showed no differences between the groups. The survival of the mice in the three different groups can be seen in Figure 5.21. One mouse from the group immunised with AP1 died 2 days post infection (Figure 5.21 A) while the other three mice all died on the 4th day post infection. The groups immunised with AP2 and the control protein (VP4) both had identical survival curves with all the mice dead by the fifth day post infection (Figure 5.21 B and C).

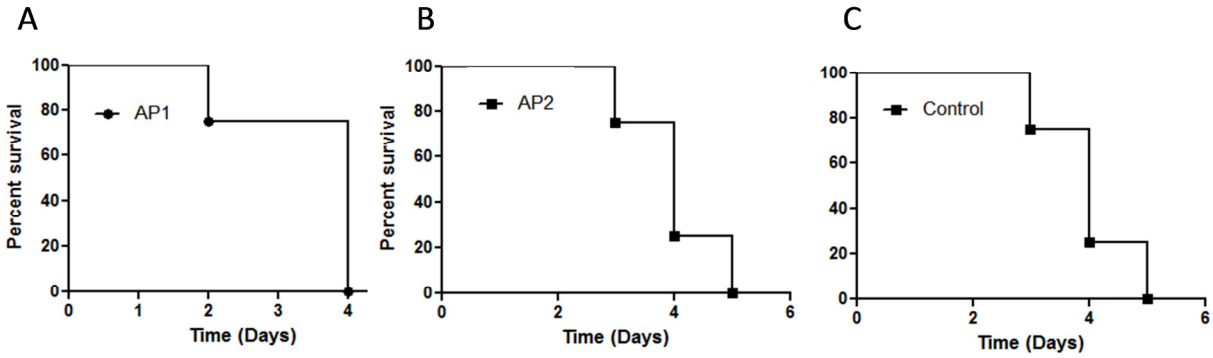


Figure 5.21 Survival of three groups of mice after infection with *T. congolense* (strain IL1180) following immunisation with AP1, AP2 or VP4 control unrelated protein. (A) group immunised with refolded recombinant AP1; (B) group immunised with refolded AP2; (C) group immunised with VP4 control antigen. Mice were immunised with 20 µg of protein in Freund's adjuvant for the initial immunisation and in Freund's incomplete adjuvant for the 3 subsequent boosters two weeks apart. Mice were rested for 2 weeks prior to infection with 1 x 10^4 *T. congolense* (strain IL1180) parasites. Parasitaemia and PCV were monitored.

5.3.6 RNAi of AP1 and AP2 in *T. congolense*

PCR was used to amplify a 350 bp region of the AP1 gene and a 400 bp region of the AP2 gene. A concatenated product of these two genes (AP1+AP2: 850 bp) was produced by a two-step PCR. These PCR products were purified prior to cloning into the p2T7^{Ti} vector for RNAi (Figure 5.22).

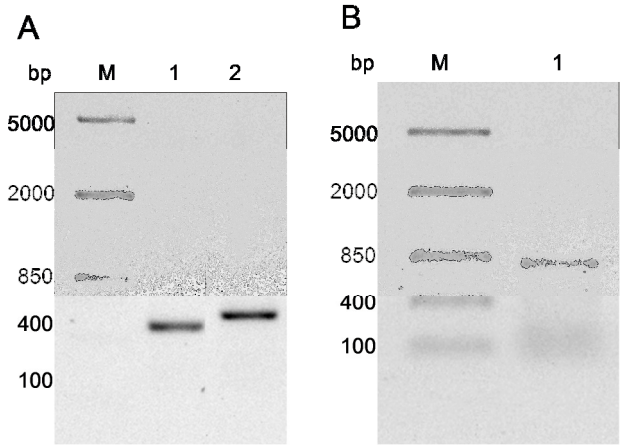


Figure 5.22 Agarose gel of purified PCR fragments for RNAi of AP1, AP2 and AP1+AP2, a concatenated product of the two. (A) Purified PCR fragments for AP1 and AP2. Lane 1: 350 bp section of AP1; lane 2: 400 bp section of AP2. (B) Purified PCR fragments for AP1+AP2 concatenated product. Lane 1: 850 bp concatenated AP1+AP2 product. M: FastRuler® DNA Ladder, Middle Range (Fermentas).

Trypanosomes (*T. congolense* strain Treu 13:29) were transfected with the prepared recombinant AP1, AP2 and AP1+AP2 - p2T7^{Ti} plasmids and grown for at least two weeks in media containing antibiotics for selection. Clones were selected (one from each construct) and genomic DNA isolated to perform a PCR on the antibiotic gene (bleomycin gene) to confirm transfection with the recombinant p2T7^{Ti} plasmids. As can be seen from Figure 5.23, a 300 bp product corresponding to a section of the bleomycin gene can be

seen in lanes 1, 2 and 3 (representing the three recombinant constructs) but not lane 4 (corresponding to the wild type untransfected strain). A non-specific PCR product of approximately 1 kb can be seen in all lanes.

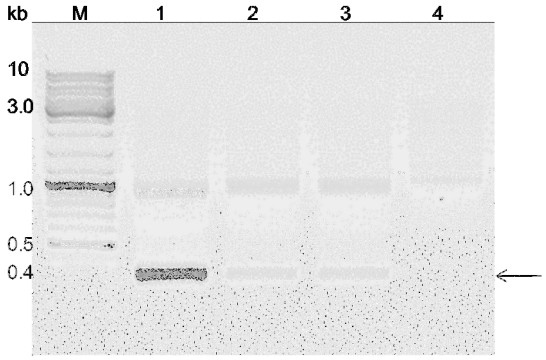


Figure 5.23 Agarose gel showing PCR to confirm insertion of the p2T7^{Ti} plasmid into *T. congolense* (strain Treu 13:29) containing the pLew13 and pLew29 vectors. PCR of a section of the bleomycin gene (p2T7^{Ti}) of 300 bp. Lane 1: AP1; lane 2:AP2; lane 3: AP1+2; lane 4: WT. Arrow indicates the 300 bp band of interest.

In order to determine whether the AP1 gene was indeed being down regulated using the tetracycline inducible construct, a western blot was performed on the *T. congolense* RNAi AP1 lysate using the anti-AP1 peptide antibody (Figure 5.24). A band of approximately 90 kDa corresponding to AP1 can be seen in lanes 1 to 7 and is absent in lanes 8 and 9. There is clearly a decrease in the amount of AP1 from days 4 to 7. Furthermore, the tubulin loading control (band at 55 kDa present in all lanes) confirms that approximately equal amounts of lysate were loaded in each of lanes 1 -7. Lanes 8 and 9 show slightly less tubulin compared to the previous lanes, therefore no conclusions can be made about the absence of AP1 from these lanes. However, since there was a clear decrease in AP1 from days 4-7 it may be surmised that the decrease was sustained for the remainder of the time analysed.

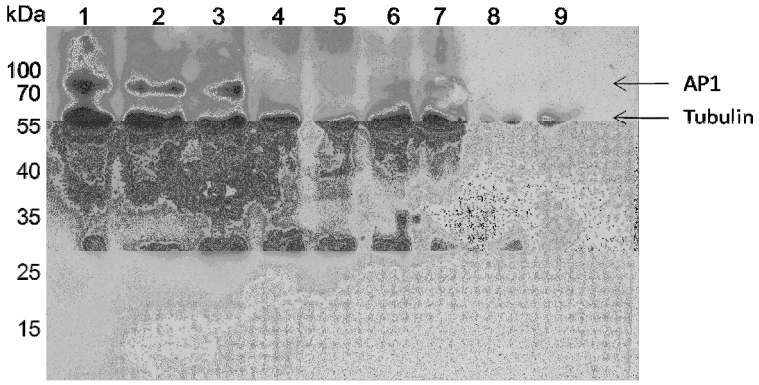


Figure 5.24 Western blot showing a decrease in AP1 expression over time in the lysate of tetracycline induced procyclic *T. congolense* p2T7^{Ti}/AP1. Lanes 1-9: lysate of *T. congolense* p2T7^{Ti}/AP1 from days 1-9 after induction with tetracycline. Initial primary antibody used was affinity purified chicken anti-AP1 IgY (1 µg/ml). The blot was re-probed (without stripping) with monoclonal mouse anti-tubulin antibody (1/2 000) as a control for loading on the SDS-PAGE. ECL immunodetection reagent was used for a substrate.

A decrease in expression of AP2 was noticed following tetracycline induction of the RNAi in the strain *T. congolense* p2T7^{Ti}/AP2 (Figure 5.25). A band at approximately 90 kDa corresponding to AP2 can be seen in lanes 1 to 6 and is absent from lanes 7 and 8. The loading control of tubulin can be seen in approximately equivalent quantities in all lanes as a triple band of approximately 55 kDa.

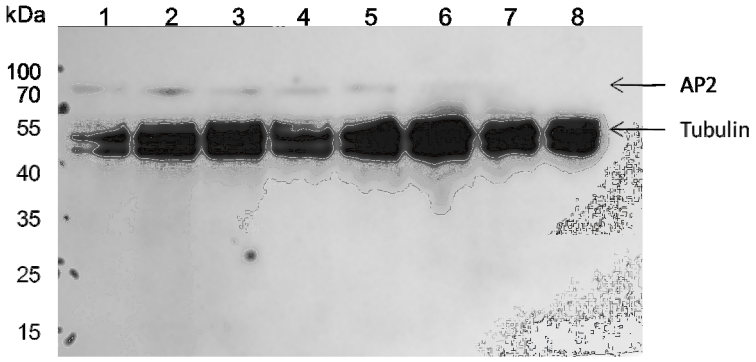


Figure 5.25 Western blot showing a decrease in AP2 expression over time in the lysate of procyclic tetracycline induced *T. congolense* p2T7^{Ti}/AP2. Lanes 1-8: lysate of *T. congolense* p2T7^{Ti}/AP2 from days 1-8 after induction with tetracycline. Initial primary antibody used was affinity purified chicken anti-AP2 peptide IgY (1 µg/ml). The blot was reprobed with monoclonal mouse anti-tubulin (1/2 000) as a control for loading of the SDS-PAGE gel. ECL immunodetection reagent was used for a substrate. Arrows indicate the bands of interest.

The wild type strain of *T. congolense* (Treu 13:29) was analysed for AP1 and AP2 expression over time. As can be seen from Figures 5.26 A and B, levels of AP1 and AP2 expression (bands in all lanes at approximately 90 kDa) remained constant over the nine days analysed.

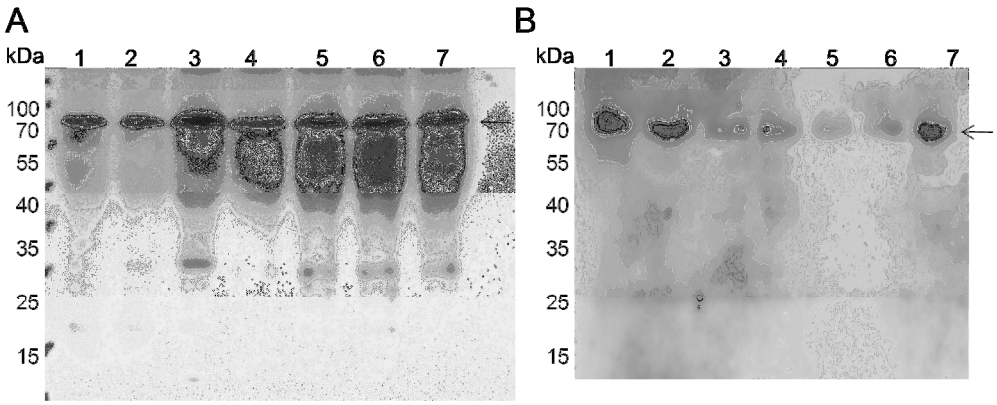


Figure 5.26 Western blot showing the expression of AP1 and AP2 over time in the lysate of procyclic wild type *T. congolense* (strain IL3000). (A and B) Lanes 1-7: lysate of *T. congolense* (strain IL3000) from days 1-7. Primary antibody used was (A) affinity purified chicken anti-AP1 peptide IgY (1 µg/ml) and (B) affinity purified chicken anti-AP2 peptide IgY (1 µg/ml). ECL immunodetection reagent was used for a substrate. The arrow indicates (A) AP1 and (B) AP2.

Reduction of AP1 and AP2 in the induced *T. congolense* p2T7^{Ti}/AP1+AP2 strain after induction with tetracycline was analysed by western blotting (Figure 5.27). As is evident from Figure 5.27 A, there is a decrease in the amount of AP1 after day 4 and it is not detectable from day 7. Similarly, Figure 5.27 B shows that there is a decrease of AP2 expression on day four and it is not detectable on day 9. The presence of equal amounts of tubulin in each lane of the blot (Figure 5.27 C) shows that equal amounts of lysate were loaded into each lane and, therefore, that the decreases in AP1 and AP2 are not a result of unequal loading.

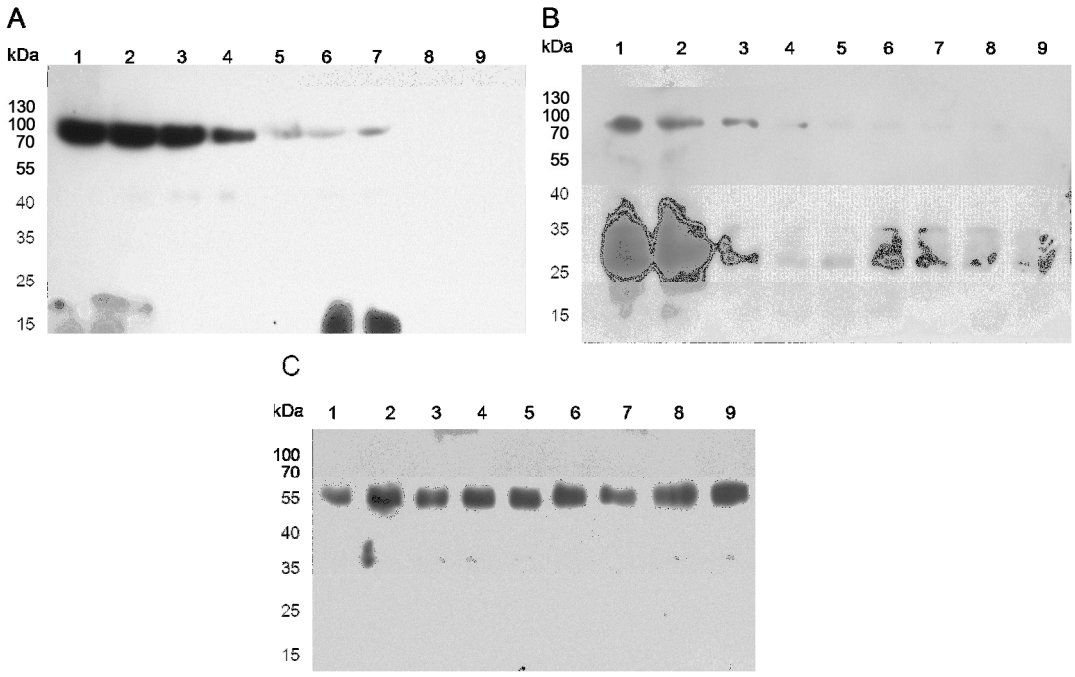


Figure 5.27 Western blot showing the decrease in AP1 and AP2 expression over time in the lysate of procyclic tetracycline induced *T. congolense* p2T7^{Ti}/AP1+AP2. (A, B and C) Lanes 1-9: lysate of *T. congolense* RNAi AP1+AP2 from days 1-9 after induction with tetracycline. Primary antibody used was (A) affinity purified chicken anti-AP1 peptide IgY (1 µg/ml), (B) affinity purified chicken anti-AP2 peptide IgY (1 µg/ml) and (C) monoclonal mouse anti-tubulin antibody (1/2000) as a control for loading of the SDS-PAGE. ECL immunodetection reagent was used for a substrate.

In addition to measuring the levels of expression of AP1 and AP2 in the RNAi induced strains by antibodies and western blotting, the activity of AP1 and AP2 were measured in lysates of the three RNAi strains, p2T7^{Ti}/AP1, p2T7^{Ti}/AP2 and p2T7^{Ti}/AP1+AP2 as well as a wild type (WT) control strain for several days after induction (Figure 5.28). It can be seen that there is a definite decrease in activity of AP1 and AP2 in all the induced strains whereas the levels of activity in the wild type strain is fairly constant. The levels of activity are at their lowest at days 8, 9 and 10, but AP activity is never completely abolished. Note that puromycin has been shown to completely inhibit activity of the alanyl APs in trypanosome lysate (see Figure 5.18).

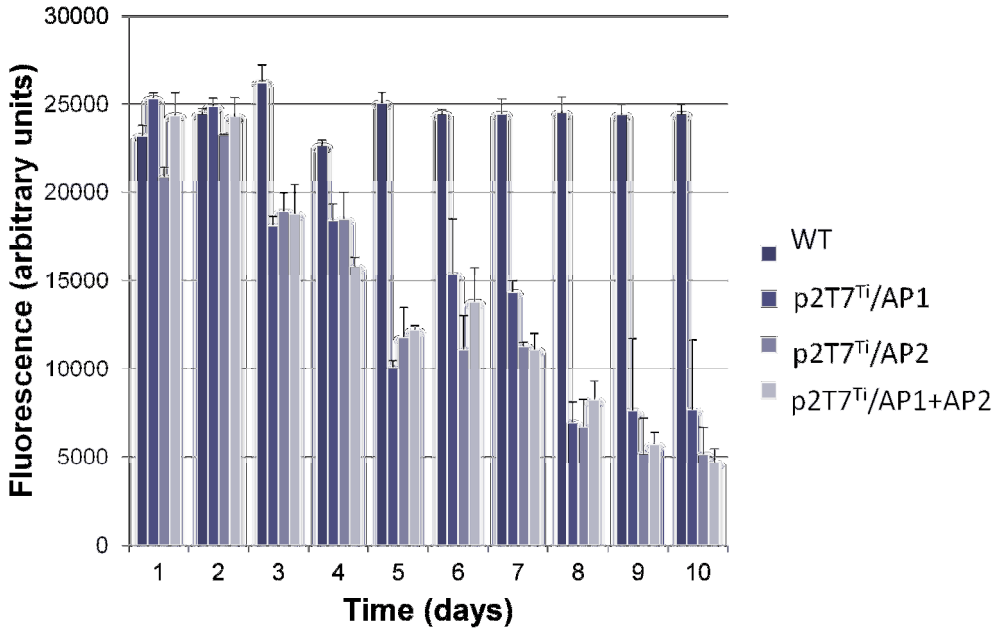


Figure 5.28 Activity assays showing a decrease in activity against H-Ala-AMC after tetracycline induction of the three RNAi AP constructs p2T7^{Ti}/AP1, p2T7^{Ti}/AP2 and p2T7^{Ti}/AP1+AP2 in procyclic *T. congolense* compared to the wild type strain. Error bars indicate the standard deviation from triplicate experiments. See Section 5.2.5 for details of method.

Growth rates of the tetracycline induced RNAi strains were compared to the wild type *T. congolense* (strain Treu 13:29) strain (Figure 5.29). Down-regulation of AP1 and AP2 by RNAi individually did not affect the growth rates of the cultures (Figure 5.29 A). In Figure 5.29 B, however, it can be seen that the induction of RNAi of both AP1 and AP2 does slightly reduce the growth rate of the trypanosomes when compared to the wild type strain.

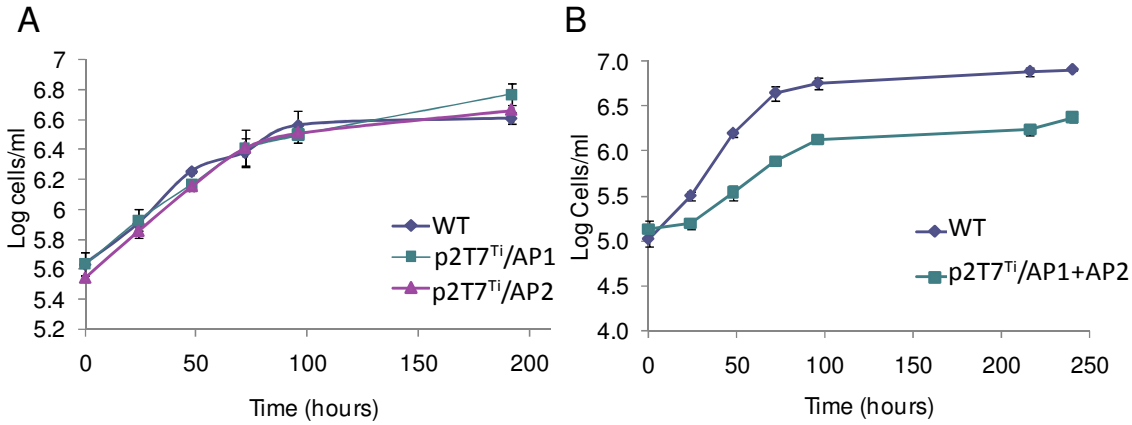


Figure 5.29 Comparison of growth of procyclic *T. congolense* (strain Treu 13:29) wild type parasites with the three tetracycline induced RNAi mutants: p2T7^{Ti}/AP1, p2T7^{Ti}/AP2 and p2T7^{Ti}/AP1+AP2. (A) Growth of the wild type, p2T7^{Ti}/AP1 and p2T7^{Ti}/AP2. (B) Growth of the wild type and p2T7^{Ti}/AP1+AP2 Cultures of each strain were diluted to approximately 5 x 10^4 cells/ml in complete MEM and grown at 27°C until saturation (10^6 – 10^7 cells/ml).

5.4 DISCUSSION

Antigenic variation has posed a problem to the development of a conventional vaccine for African animal trypanosomosis. For this reason the idea of an anti-disease vaccine, whereby the disease causing factors of the parasite are targeted rather than the parasite itself, has been proposed (Authié et al., 2001; Antoine-Moussiaux et al., 2008). The M1 APs of *T. congolense* have been identified as possible pathogenic factors (see Section 2.3), that is, parasite antigens responsible for the development of disease in the host. Although several APs from parasites such as *Plasmodium falciparum* (Skinner-Adams *et al.*, 2009) and *Fasciola hepatica (Acosta et al., 2008)* have been evaluated as potential vaccine candidates or drug targets, APs have never before been investigated in any trypanosome species. For these reasons, this study has been conducted to identify and characterise two M1 type APs from *T. congolense,* named AP1 and AP2.

In the present study, AP1 and AP2 were cloned, sequenced, and expressed. Initially, expression of AP1 and AP2 was attempted in the eukaryotic system, *Pichia pastoris*, since the APs had a high number of predicted disulfide bridges (11 and 6 respectively, data not shown) and would be likely to undergo post-translational modifications similar to that of the native proteins resulting in solubly expressed functional proteins (Cereghino *et al.*, 2002). However, even after altering media, amount of methanol for induction, vectors and strains of *P. pastoris,* no expression was achieved (data not shown). Expression was subsequently conducted in *E. coli* (BL21 DE3) and, as expected, the recombinant proteins were expressed as insoluble inclusion bodies using the pGEX-4T-1, pET28a and pET32a vectors. It was necessary to first solubilise and, thereafter, refold the expressed APs to obtain functional enzymes.

It is possible, although uncommon, for inclusion bodies to be similar to the native form of the protein and easy to solubilise using non-ionic detergents or low concentrations of sodium chloride (Burgess, 2009). However, this is rarely the case as the difficulty of solubilising AP1 and AP2 inclusion bodies in 8 M guanidine demonstrated. Given the high number of cysteine residues in both APs, a reducing agent (DTT) was added to the inclusion body solubilisation buffer (but omitted from the refolding buffer) in order to minimise incorrect formation of disulfide bridges (Burgess, 2009). Following solubilisation of the inclusion bodies, a refolding kit was used, facilitating the parallel testing of several different refolding buffers. Each different refolding buffer comprised the same base buffer (see Section 5.3.1) to which varying quantities of guanidine (used to solubilise the inclusion bodies), L-arginine, and ratio of oxidised (GSSG) to reduced (GSH) glutathione

were added. A similar, but much more complex, 96 well system was employed by Sijwali *et al.* (2002) who refolded falcipain-2, a cysteine peptidase of *P. falciparum* using this systematic approach. Falcipain-2, with 4 predicted disulfide bridges, was successfully refolded and autocatalytically processed using a 100-fold dilution into a buffer containing 30% glycerol, 250 mM L-arginine and 1 mM each of GSH and GSSG at alkaline pH (Sijwali *et al.*, 2001).

This refolding process is dependent on significant dilution of the inclusion bodies, usually to 10-100 µg/ml, (Burgess, 2009) as protein concentration is the chief factor in solubilisation and lower concentrations of protein favour refolding as opposed to aggregation (Singh and Panda, 2005). It is possible to refold proteins at higher concentrations (up to 5 mg/ml) by adding a low concentration of a chaotrope such as guanidine or urea, however, this approach is not often successful (Menzella *et al.*, 2002). The addition of L-arginine reduces aggregation of mis-folded proteins by slowing the rate of protein-protein interactions thereby increasing the solubilisation of the protein (Lange and Rudolph, 2009). One disadvantage of using L-arginine is that it is most effective at relatively high concentrations (0.5 – 1 M) but this level of L-arginine interferes with downstream protein purification methods including nickel chelate affinity- and ion exchange chromatography, thereby necessitating a dialysis or dilution step (Cabrita and Bottomley, 2004).

Following solubilisation, reforming of disulfide bonds has to be promoted in order to obtain correctly folded protein, and this was achieved by the addition of the glutathione redox couple. Disulfide bonds in proteins are formed by thiol-disulfide exchange whereby the thiolate group ($-S^-$) of a cysteine residue attacks the sulfur atom of a disulfide bond (provided by oxidised glutathione, GSSG, in the present case) (Figure 5.30). The original disulfide bond is broken by the attachment of the thiolate group and the sulfur atom is transferred to the attacking thiolate resulting in the formation of a mixed disulfide. Attack of this disulfide bond by a second protein thiolate (provided by a cysteine residue in close proximity) promotes formation of an intra-molecular disulfide bond. Reduced glutathione (GSH) is responsible for reducing the mixed disulfides which form initially and thereby promoting intra-molecular disulfide formation (Bulaj, 2005). GSH is subsequently oxidised to GSSG which in turn is converted to a protein thiolate, and the sequence of reactions is repeated for a subsequent disulfide bond. For this reason, the initial reaction requires a low concentration of GSSG but higher concentration of GSH. Optimising the molar ratios

of the two forms is, therefore, essential for the correct formation of intra-molecular disulfide bonds during refolding of a protein.

Figure 5.30 Mechanism of intra-molecular disulfide bond formation (Bulaj, 2005). In the present study the RSSR represents oxidised glutathione (GSSG) which in turn is converted to reduced glutathione (GSH) when a thiolate ($-S^-$) of the protein attacks the disulfide bridge. This newly formed disulfide (protein-SSG) can then be attacked in a similar fashion by a nearby thiolate on the protein resulting in the formation of GSH and an intramolecular disulfide bond.

The efficiency of the refolding of APs was evaluated in three ways: firstly, by testing cleavage of the synthetic substrate H-Ala-AMC indicating that the APs were in an active conformation; secondly by visually observing the presence or absence of a precipitate in the solution after refolding, indicating insoluble protein; and lastly, by a difference in size between reduced and non-reduced protein on SDS-PAGE. As mentioned, the problem with using the presence of a precipitate to determine the level of refolding is that the samples used for refolding were partially purified inclusion bodies, and not homogeneous solutions of protein. Therefore, it is possible that the precipitated protein could be a contaminating protein and not the protein of interest. This is likely to be the case for AP2, which had the highest level of activity in Buffer 1 but also the greatest amount of precipitate (see Table 5.1).

Evaluating the shift in size between reduced and non-reduced samples on SDS-PAGE gel was also not a very useful method to determine the relative efficiencies of the different refolding buffers. This is due to the fact that the samples had to be very dilute for the refolding process and consequently were fairly difficult to visualise by SDS-PAGE, thus making it difficult to see relative differences between the effects of the different refolding buffers (see Figure 5.8). As a result, the preferred buffers for refolding of the two APs were chosen based both on the level of activity against H-Ala-AMC and whether there was a visible shift in size of reduced vs. non-reduced samples on SDS-PAGE. Following refolding, the two APs were purified to homogeneity by His-affinity chromatography followed by ion exchange chromatography.

The catalytic activity of the purified APs was tested on several single amino acid synthetic substrates. AP1 and AP2 had quite similar catalytic properties, including comparable K_m values for all four substrates that were successfully hydrolysed. As expected, the AP1 and AP2 hydrolysed H-Ala-AMC most efficiently, with a specificity constant for AP1 almost four fold higher than that of the next most preferred substrate, H-Gly-AMC. Furthermore, both APs did not cleave H-Arg-AMC or H-Tyr-AMC. It can be concluded that AP1 and AP2 are indeed alanyl aminopeptidases belonging to the M1 Family with a very defined specificity for small non-polar R-groups, most preferably alanine (Taylor, 1993).

Like the M17 APs of the parasites *P. falciparum* (Stack *et al.*, 2007), *L. amazonensis* (Morty and Morehead, 2002) and *Setaria cervi* (Pokharel and Rathaur, 2008), activity of the two APs of *T. congolense* was inhibited by the addition of Cu^{2+}, but increased by the addition of Mg^{2+}, Mn^{2+}, Fe^{2+} , Ni^{2+} and Co^{2+}. The neutral pH optima for AP1 and AP2 (8.0 and 7.5 respectively) also confirm that the aminopeptidases are likely to be cytosolic (further confirmed by the immunofluorescence localisation in the parasite) and not confined to the endocytic system like many other peptidases (Rawlings and Barrett, 1993). Furthermore, inhibition of both AP1 and AP2 by the antibiotic puromycin distinguishes the two APs from membrane bound APs also belonging to Family M1 (Gros *et al.*, 1985).

Determining the stages of the lifecycle at which the APs were expressed is essential if these peptidases are to be evaluated as drug targets or vaccine candidates. It was found that AP1 and AP2 were expressed at comparable levels throughout the life-cycle of the parasite. This is also the case for the M17 aminopeptidases of the parasites *Fasciola hepatica* (Acosta *et al.*, 2007), *L. amazonensis* (Morty and Morehead, 2002), *P. falciparum* (Stack *et al.*, 2007) and *S. cervi* (Pokharel and Rathaur, 2008). However, it has been postulated that these APs are more important to the parasite in the life cycle stage in the mammalian host especially for the intracellular parasites such as *P. falciparum* and the *Leishmania* spp. since the APs are solely responsible for the liberation of individual amino acids necessary for protein synthesis (Skinner-Adams *et al.*, 2009).

In an attempt to elucidate the function of the APs in the parasite, RNAi was used to down-regulate AP expression *in vitro*. First coherently described by subsequent Nobel winners Fire and Mello in *Caenorhabditis elegans* (Fire *et al.*, 1998), RNAi is an intrinsic cellular mechanism found in some eukaryotes in which the presence of double stranded RNA causes mRNA with the same sequence to be degraded, effectively silencing expression of

that gene (LaCount and Donelson, 2001). RNAi was first successfully employed in trypanosomes by Ngo *et al.* (1998) who transiently down-regulated α-tubulin expression in *T. brucei* producing the so-called "FAT" phenotype which is still regularly used as a control for RNAi in trypanosomes (Coustou *et al.*, 2010). Subsequent improvements of the original method include regulation of RNAi by use of a tetracycline inducible system (Wirtz *et al.*, 1999); stable, heritable transfection as opposed to transient transfection (Bastin *et al.*, 2000a; Shi *et al.*, 2000; Wang *et al.*, 2000); and adaptation of the entire system to *T. congolense* (Inoue *et al.*, 2002). A comprehensive review of RNA interference in protozoan parasites has been published by Ullu *et al.* (2004).

Some of the benefits of RNAi include the fact that it is relatively quick and easy to perform when compared to gene knock-out. Also, RNAi of two or more different genes is possible by cloning concatenated sections of genes, whereas knock-out of multiple genes requires successive transfections which is difficult in *T. congolense*, if not impossible. Similarly, diploid organisms do not require successive transfections for down-regulation using RNAi as would be necessary for gene knock-out (Bellofatto and Palenchar, 2008). Possible disadvantages of RNAi include what is known as "off target effects", where the introduced RNAi interferes with the expression of genes other than the target gene. This can be fairly easily avoided by cross-checking the section of RNA to be targeted against other genes in the trypanosome genome database. Another possible disadvantage is the leakage of expression of the double stranded RNA without the addition of tetracycline. This background expression can be especially problematic if the targeted gene is essential. If this is not the case, most problems associated with background expression can be minimised by using the wild type strain (rather than simply a non-induced transgenic strain) as a control. A further issue with RNAi is that the success of the down-regulation is dependent on the stability of the targeted RNA i.e. rapidly transcribed, long-lived mRNA would be more difficult to down-regulate than unstable mRNA expressed at low levels (Bellofatto and Palenchar, 2008).

It was found that decreased expression of each of the APs individually had no observable effect on the *T. congolense* phenotype. However, down-regulation of both the APs simultaneously appeared to slightly inhibit the growth rate of procyclic *T. congolense*. Furthermore, it proved impossible to adapt these parasites back to *in vitro* culture after multiplication of the bloodstream form in mice. It is possible that the decrease in AP expression affects growth of the parasite more as bloodstream forms than as procyclics, therefore preventing growth of the recombinant parasites. RNAi of APs in the parasitic

nematode *Haemonchus contortus* (Samarasinghe *et al.*, 2010), the helminth parasite *Schistosoma mansoni* (Rinaldi *et al.*, 2009) and the hard tick *Haemaphysalis longicornis* (Hatta *et al.*, 2007) have shown that the APs reduce parasite viability by decreasing reproductive potential and limiting egg output. It is postulated that a lack of APs reduces the free amino acids necessary for novel protein synthesis thereby reducing fecundity (Samarasinghe *et al.*, 2010).

RNAi down-regulation of peptidases of several parasites has shown promising results. Firstly, the RNAi-induced decrease in *T. b. brucei* cathepsin B prevented death and cured mice infected with *T. b. brucei* (strain 90-13), which is usually lethal (Abdulla *et al.*, 2008). Additionally, down-regulation of brucipain (*T. b. brucei* cathepsin L) prolonged survival in infected mice, as well as decreased invasion of an *in vitro* model of the human blood-brain barrier (Abdulla *et al.*, 2008). RNAi of cathepsin D of *S. mansoni* resulted in decreased growth *in vitro* and was fatal to the parasites *in vivo*, probably due to the inability of the cathepsin D deficient parasites to degrade haemoglobin (Morales *et al.*, 2008). Interestingly, RNAi mediated suppression of cathepsin B from *S. mansoni* did not influence the parasites' ability to degrade haemoglobin but did result in growth retardation *in vitro* and *in vivo* (Correnti *et al.*, 2005). Protozoa of the genus *Acanthamoeba* are opportunistic parasites of varying levels of pathogenicity causing diseases such as granulomatous encephalitis, which can be lethal, and eye keratitis. For some strains of *Acanthamoeba*, down-regulation of a secreted serine peptidase resulted in a decrease in cytotoxicity, implying a role for this peptidase in pathogenicity and virulence (Lorenzo-Morales *et al.*, 2005).

It was not unexpected that immunisation with the APs individually would not provide protection to mice upon challenge with *T. congolense* since it is evident that the two APs are not the only pathogenic factors of *T. congolense.* Immunisation of susceptible cattle with two isoforms of congopain, known to be a pathogenic factor, did not result in complete protection against disease. However, the immunised cattle recovered PCV two to three months after the initial acute onset of anaemia (Authié *et al.*, 2001). Immunisation of mice with oligopeptidase B, also a pathogenic factor of *T. congolense* (Coetzer *et al.*, 2008), did not prevent death but did prolong survival by 40% upon challenge with *T. congolense* (strain IL1180) (Bizaaré, 2008).

Grebaut *et al.* (2009) attempted to identify novel pathogenic factors by examining the variable expression of excreted/secreted proteins between *T. congolense* strains of

differing virulence and pathogenicity, and have identified a large group of proteins including glycolytic enzymes, cytoskeletal proteins, cysteine peptidases and an M17 type aminopeptidase. For an anti-disease vaccine to be protective, it would have to include all pathogenic factors, including the two APs described in the present study.

CHAPTER 6: GENERAL DISCUSSION

African animal trypanosomosis (AAT) is a wasting disease of livestock resulting in economic losses estimated at USD $ 1 – 5 billion per annum in Africa (Kristjanson *et al.*, 1999). Current control methods for AAT are insufficient and there is no vaccine for AAT. In an attempt to circumvent the problem posed by antigenic variation, the idea of an anti-disease vaccine has been proposed for AAT (Authié et al., 1993b; Serveau et al., 2003; Antoine-Moussiaux et al., 2009).

An anti-disease vaccine is designed around the concept that pathogenicity is caused by compounds released by the parasites (pathogenic factors) rather than the parasite itself (Schofield, 2007). The phenomenon of trypanotolerance, whereby some species of breeds of mammals are able to tolerate parasites in their bloodstream without severe pathogenesis, is strong evidence for the feasibility of an anti-disease vaccine (Naessens *et al.*, 2002). For an anti-disease vaccine to be successful, it has to include all the factors which contribute to pathogenesis. Therefore, an essential step in the development of an anti-disease vaccine is the identification of these pathogenic factors. In the present study, two proteomic tools have been used to identify these pathogenic factors in *T. congolense* with varying levels of success. The use of an affinity column permitting selective identification of antigens specifically recognised by trypanotolerant cattle has allowed the discovery of several potential pathogenic factors. Based upon current studies in other protozoan parasites (Acosta et al., 2008; Marcilla et al., 2008; Cazorla et al., 2009; Duschak and Couto, 2009), a group of cathepsin L-like CPs and two M1 type aminopeptidases were selected for further study.

Congopain is a cathepsin L-like CP belonging to Clan CA of the CP family and, therefore, has a classical papain-like catalytic triad of Cys^{25}, His^{159} and Asn^{175} (papain numbering). The pro-peptide domain (120 residues) is auto-catalytically cleaved at low pH (Serveau *et al.*, 2003) to release the mature enzyme consisting of a catalytic domain connected to a C-terminal domain by a poly-proline hinge region. The C-terminal region is unique to parasite cysteine peptidases and completely absent in the mammalian counterparts (Mottram *et al.*, 1989). There is currently no concrete evidence as to the function of this C-terminal region. It has, however, been suggested that the C-terminal extension may play a role in targeting the enzyme to the lysosome or glycosome, as is the case for the

phosphoglycerate kinase of *Crithidia fasciculata*, another kinetoplastid organism (Swinkels *et al*., 1988). Additionally, there is a predicted N-glycosylation site in the C-terminal region of congopain which could potentially allow attachment of a lysosome targeting marker.

Congopain has been enzymatically characterised (Mbawa et al., 1992; Authié et al., 1993a; Chagas et al., 1997) and is now known to be active against conventional synthetic CP substrates over a broad pH range (after the autocatalytic cleavage of the pro-peptide at acidic pH). Like most papain family peptidases, congopain has endoproteolytic activity (Mbawa *et al*., 1992) and preferentially binds bulky, hydrophobic residues (such as phenylalanine) in P_2 (Lecaille *et al*., 2001a). However, the only *in vivo* substrates of congopain identified so far are the variant surface glycoproteins which are thought to be intracellularly processed by parasite CPs (Mbawa *et al*., 1992). Upon selection of a cDNA encoding a *T. congolense* CP (Fish *et al*., 1995, EMBL accession Z25813, CP1), it was presumed at the time that the CP gene encoded congopain. It was later established, once a cDNA coding for a second CP that translated into congopain was found (Jaye *et al*., 1994 EMBL accession L25130, CP2), that the initial cDNA coded for a different CP (now called CP1), and not congopain based on N-terminal sequencing. This was the first evidence that multiple genes coding for cathepsin L-like CPs are present in *T. congolense.* However, immunisation using CP1 and congopain has been shown to provide only partial protection against trypanosomosis (Authié *et al*., 2001).

It has since been shown by Southern blotting that the cathepsin L-like CPs of *T. congolense*, like those of *T. brucei* (Mottram *et al*., 1989), *T. cruzi* (Eakin *et al*., 1992) and *L. mexicana* (Souza *et al*., 1992) are present in a tandem array consisting of at least 13 genes organised in 2 kb gene units (Kakundi, 2008). The presence of multiple congopain-like genes might be the explanation as to why only partial protection was found after immunisation with CP1 and congopain (Authié *et al*., 2001). The present study was aimed at investigating the level of variability among the multi-copy congopain-like genes and whether these multiple genes code for functional peptidases which are significantly different from congopain in terms of substrate specificity or inhibitor profiles.

Given that the *T. congolense* genome has not yet been completely sequenced and that the existing data are extremely preliminary, the sequences of the genes in the congopain locus had to be obtained from DNA libraries for the present study. Digestion of fifteen cosmids encompassing the congopain gene locus produced a major band at 2 kb (established as the size of the congopain gene units) that was cloned to constitute the

initial library. A PCR-based approach was also attempted in which degenerate primers based on an alignment of 29 database reads were used to amplify whole gene units from *T. congolense* genomic DNA. It was thought that these two approaches, together with database mining, would identify all the congopain-like CPs present in the *T. congolense* genome.

One of the most important observations which can be made from these collections of sequences is that the YHNGA pro-peptide motif (residues 127 – 132) found in congopain is highly conserved. For mammalian CPs, the pro-peptide is involved in the regulation of peptidase activity by blocking the active site, thus preventing the binding and, therefore, catalysis, of the substrate (Cygler and Mort, 1997). This phenomenon has been exploited in the development of cathepsin inhibitors designed based on the pro-peptide of the peptidase to be inhibited (Carmona et al., 1996; Billington et al., 2000; Guay et al., 2000; Reis et al., 2007). These pro-peptide-derived inhibitors have shown a limited specificity, as those inhibitors designed based on the pro-peptide of a cathepsin L family peptidase (cathepsins L, K and S) did not inhibit cathepsin B or papain (Guay *et al.*, 2000). Similarly, the pro-peptide sequence of trypanosomal CPs was used to design inhibitors which regulate the specific enzyme activity (Lalmanach et al., 1998; Godat et al., 2005). Lalmanach *et al.* (1998) showed that a peptide containing the YHNGA motif inhibited congopain and cruzain but not mammalian cathepsins B or L. In a later study, Godat *et al*. (2005) showed that this inhibition could be enhanced by the addition of part of a transcription factor from *Drosophila antennepedia* to the congopain pro-peptide-derived inhibitor. More recently, an inhibitor derived from the pro-peptide of cruzain was shown to inhibit the trypanosomal CPs, cruzain, brucipain and human cathepsin L, but not human cathepsins S, K or B (Reis *et al.*, 2007). Therefore, the fact that the pro-peptide is highly conserved in the congopain-like peptidases is viewed to be significant as it could mean that a single inhibitor could potentially be used against all of the peptidases encoded by the congopain-like genes identified from the libraries constructed in the present study.

Chemotherapeutic drugs targeting cysteine peptidases have been explored for several parasitic diseases including malaria (Rosenthal, 2004; Ettari et al., 2010), Chagas disease (Duschak and Couto, 2009; Brak et al., 2010), leishmaniasis (Bryson *et al.*, 2009) and fasciolosis in sheep (Roche et al., 1999; Alcala-Canto et al., 2007). Interestingly, the leishmanial cathepsin L-like enzymes (encoded by the *lmcpb* array), although not needed for the survival of the parasite, have been shown to be necessary for virulence i.e. as pathogenic factors (Mottram et al., 1996; Denise et al., 2003). Therefore, inhibition of

these peptidases would not necessarily result in death of the parasite, but could reduce virulence. However, the inhibition of cruzain, the major cathepsin L-like cysteine peptidase of *T. cruzi*, has been shown to indirectly cause the death of the parasite by blocking the normal processing and trafficking of cruzain resulting in major abnormalities of the Golgi complex (Engel *et al.*, 1998). No knock-out studies of the cathepsin L-like cysteine peptidases of *T. congolense* or *T. brucei* have been published yet, so it is still uncertain whether congopain, or any other similar enzymes, are essential for parasite survival. It has, however, been demonstrated that cysteine peptidase inhibitors which interact with the major cysteine peptidase of *T. b. brucei* result in the death of the parasite *in vitro* and cure mice from an otherwise lethal dose of parasites, although the exact mechanism is unknown (Troeberg et al., 1999; Troeberg et al., 2000).

From the three different sources of congopain-like CP genes explored in this study 22, 24 and 29 unique CP genes were identified, which are all significantly more than the initial estimate of 13-14 genes obtained from Southern blotting (Boulangé, unpublished). This could possibly be explained by the fact that trypanosomes are diploid i.e. a single locus of 13-14 genes exists on each homologous chromosome resulting in 29 genes that are dissimilar due to allelic differences. To put these numbers into context, there are at least twenty cathepsin L-like peptidases in *T. brucei* which are found in a long tandem array consisting of gene units of approximately 1.75 kb (Mottram *et al.*, 1989). *T. cruzi* is estimated to have 14 and 23 copies of cruzain in two allelic clusters found on homologous chromosomes (Tomás and Kelly, 1996). Interestingly, the last gene (3') in the cruzain array is truncated at the end of the region coding for the catalytic domain and, therefore, codes for a cruzain-like enzyme without a C-terminal extension (Tomás and Kelly, 1996). Similarly, the *lmcpb* array encoding the cathepsin L-like genes of *L. mexicana* contains 19 genes with 97–99 % sequence identity (Mottram *et al.*, 1996).

It was hypothesised that the reason for expression of multiple forms of similar enzymes is that each isoform has unique substrate preferences corresponding to a unique role in the parasite (Boulangé *et al.*, 2001). Moreover, there is evidence that the multiple cathepsin L-like genes encode functionally distinct stage-regulated enzymes in both *L. mexicana* (Mottram *et al.*, 1997) and *T. cruzi* (Lima *et al.*, 2001). In order to determine whether this was indeed the case with the congopain-like genes of *T. congolense*, expression of variant truncated catalytic domain congopain-like peptidases was conducted. Kinetic experiments performed in the present study with the recombinant peptidases suggest that the differences in primary sequence do indeed result in differences in substrate

specificities of the different peptidases. However, all three expressed peptidases cleaved substrates with a hydrophobic amino acid residue in P_2 and amino acid residues with a polar side chain in P_1. Substrates with a hydrophobic amino acid residue in P_1 were not hydrolysed at all. Given the slight differences in substrate preferences, the results obtained in the present study in terms of enzyme kinetics are insufficient to conclude whether the different congopain-like peptidases have distinct roles in the parasite. Besides conducting more extensive enzyme substrate kinetics assays, gene disruption technology could be used to determine the role of the individual congopain-like peptidases in the parasite as was done in *L. mexicana* (Mottram *et al*., 1997). For the study in *L. mexicana*, the entire *lmcpb* array encoding the cathepsin L-like genes was knocked out resulting in parasites of reduced virulence. By subsequent re-expression of individual genes and examination of the resulting phenotype, the function of the protein encoded by individual genes could be determined (Denise *et al*., 2003).

One interesting finding of the present study is the fact that a gene encoding an inactive (CP_{SSN}) congopain-like peptidase is present in the congopain array. This was thought to be unusual since the specificity of an enzyme, and its ability to function correctly is highly sensitive to changes in the amino acid residues involved in catalysis. Hence it is presumed that these residues would be highly conserved. Any mutation in the gene resulting in a change to these catalytic residues causing a loss in specificity or total activity of the enzyme would most likely result in the gene being lost from the organism over time or during cell division. However, the presence of genes encoding inactive enzymes may not be as rare as first thought. In fact, as a more recent study has stated: “inactive enzyme-homologues are the rule, not the exception” (Pils and Schultz, 2004b). This was based upon the finding (among others) that in the seven species investigated (including *Drosophila melanogaster, Mus musculus* and *Homo sapiens*) at least 10% of the papain family cysteine peptidase genes present in the respective genomes code for inactive peptidases. Therefore, the fact that genes encoding enzymes with unusual catalytic residues are still present in the genomes of diverse organisms, may indicate that the products of these genes still have a functional role in the organism (Pils and Schultz, 2004b). One possibility is that the inactive enzymes regulate functioning of the active enzymes by binding to substrate and effectively removing it from the reaction as has been shown for the protein tyrosine phosphatases (Pils and Schultz, 2004a) .

An additional noteworthy discovery made in the present study was that a cysteine peptidase with a serine substitution for the catalytic cysteine (CD_{SHN}) is active against

classical cysteine peptidase substrates and is inhibited by the classical cysteine peptidase inhibitor E-64. The presence of multiple genes encoding peptidases with unusual catalytic sites is not unique to the congopain-like peptidases of *T. congolense.* For example, the serine repeat antigen 5 (SERA5) is a protein expressed by *Plasmodium falciparum* and is one of a nine gene family which are differentially expressed throughout the lifecycle of the parasite (Putrianti *et al.*, 2010). The eight other SERA proteins are characterised by a papain-like domain. In SERA5, however, the catalytic cysteine residue is replaced by a serine. This substitution seems to have resulted in what is, by sequence homology, a cysteine peptidase but possessing serine peptidase characteristics including inhibition by the classical serine peptidase inhibitor 3, 4 diisocoumarin (Hodder *et al.*, 2003). This was not the case for the peptidase investigated in the present study. The inverse case is found with the poliovirus peptidase 3C, which resembles a serine peptidase (specifically chymotrypsin) in terms of three-dimensional structure, but possesses a catalytic cysteine residue instead of a serine. Interestingly, active site mutation of the catalytic cysteine residue to serine resulted in a slight decrease in activity (Sarkany and Polgar, 2003).

A further consideration is that, even with differences in primary structure and substrate preferences, all three expressed variants of congopain share a pH optimum of 6.0-6.5. Congopain has previously been localised to the lysosomal system of the parasite, which correlates with the peptidase's probable role in intracellular proteolytic processing (Mbawa *et al.*, 1991b). It is, therefore, quite surprising that the peak of congopain activity is at an almost neutral pH, given that the pH of the lysosomal system is normally acidic (pH 4.2) (Mbawa *et al.*, 1992). The fact that these different enzymes are all functional at physiological pH is strong evidence for the idea that these enzymes are indeed expressed by the parasite and have distinct roles *in vivo* other than that of the intracellular proteolytic processing taking place in the lysosome.

As mentioned previously, the roles of peptidases encoded by individual congopain-like genes could be elucidated using the knock-out and individual knock-in technique developed by Mottram *et al.* (1997) for use in *L. mexicana*. However, due to the extremely high sequence identity between the intergenic regions between each of the genes on the multiple gene array of CPs in *T. congolense,* it would be exceedingly difficult to delete or insert individual genes.

In addition to the *T. congolense* cathepsin L-like CPs, the Family M1 aminopeptidases (APs) were also identified as potential pathogenic factors. The Family M1 APs are

metallopeptidases requiring a divalent metal co-factor for the catalysis of single amino acid residues from the N-terminal end of peptides. Aminopeptidases have never before been studied in any trypanosome species. The leucyl APs of *F. hepatica* and *H. contortus* have been evaluated for their protective potential in conventional vaccines with some success (Acosta et al., 2008). Interestingly, the leucyl AP of *F. hepatica* was used in conjunction with the major cysteine peptidase as a cocktail vaccine. However, best protection (89%) was achieved by use of the leucyl AP alone (Piacenza *et al*., 1999). Furthermore it has been shown that inhibition of the M1 APs of *P. falciparum* results in growth inhibition *in vitro* (Gardiner *et al*., 2009; Skinner-Adams *et al*., 2009).

Two APs of *T. congolense*, AP1 and AP2 were identified in the present study as potential pathogenic factors using immunoprecipitation. These APs were successfully expressed as insoluble inclusion bodies in an *E. coli* system after attempts at expression in the eukaryotic *Pichia pastoris* system failed. Several similar APs have previously been expressed as soluble active proteins in *E. coli* including the leucyl APs of *F. hepatica* and those of three *Leishmania* species, all of which had an added histidine tag to aid purification (Morty and Morehead, 2002; Acosta et al., 2007). Similarly, the methionine AP of *P. falciparum;* the leucyl APs of the tick *Haemaphysalis longicornis;* apicomplexan parasite *Toxoplasma gondii* and the bacterium *Salmonella enterica,* were all expressed as a soluble active enzymes using a GST fusion protein tag in an *E. coli* expression system (Mathew et al., 2000; Chen et al., 2006; Hatta et al., 2006; Jia et al., 2010). The leucyl APs of the flukes *Schistosoma mansoni* and *S. japonicum*, however, were expressed as soluble functional enzymes in a eukaryotic baculovirus expression system (McCarthy *et al*., 2004). It is unclear exactly why expression of the *T. congolense* APs did not succeed in the yeast expression system. The only report of successful expression of an aminopeptidases in *P. pastoris* is that of the pathogenic fungus *Trichophyton rubrum* (Monod *et al*., 2005). It is possible that the expression of active APs may be toxic to the yeast, as was the case for bovine trypsinogen, the zymogen of the serine peptidase trypsin. Processing of the zymogen during expression resulted in the active form of trypsin being released and very little recombinant protein being recovered. It was found that preventing this maturation by a single point mutation resulted in large amounts of the inactive unprocessed zymogen accumulating during expression in *P. pastoris* (Hanquier *et al*., 2003).

The inclusion bodies containing the recombinant APs were solubilised using high levels of guanidine and a screening process to determine the optimum conditions for refolding was

conducted. This systematic approach involved changing the amount of L-arginine, guanidine, and ratio of oxidised to reduced glutathione to determine optimum concentrations of each reagent required for refolding. The only report of another AP to be refolded is that of porcine kidney leucyl AP (Laslo *et al.*, 2009). This AP, however, was refolded by use of "artificial chaperones", a technique pioneered by researchers Rozema and Gellman (1996). The principle of the method is as follows: the denatured protein is mixed with the cationic detergent cetyltrimethylammonium bromide which prevents aggregation by binding to the protein forming a complex. The subsequent removal of this detergent using a polysaccharide (dextrin-10 in the referenced study) allows spontaneous refolding of the protein. This method was highly successful for the refolding of porcine leucyl AP, producing 92% correctly folded protein (Laslo *et al.*, 2009).

In the present study, the success of refolding was assessed by three approaches. Firstly the presence of a precipitate in the refolding solution could indicate that the protein was incorrectly folded and therefore precipitating out of solution. However, since the refolding was not conducted using pure protein, but rather purified inclusion bodies (containing approximately 80% of the protein of interest); it is possible that some of the precipitate observed were contaminating proteins. A second approach for assessing refolding was to observe the differential migration of the reduced (by addition of β-mercaptoethanol and boiling) and non-reduced forms of the protein after incubation in the refolding buffer. Non-reduced, refolded protein migrated faster than reduced, refolded protein since all intra-molecular disulfide bonds were intact, rendering the protein more compact. Lastly, since the APs are enzymes, correctly refolded AP1 and AP2 were shown to be active and catalysed the hydrolysis of the synthetic substrate H-Ala-AMC.

The kinetic characteristics of the refolded AP1 and AP2 enzymes were determined. It was found that the two enzymes did indeed preferentially cleave the substrate H-Ala-AMC better than H-Leu-AMC, H-Gly-AMC and H-Met-AMC, indicating that these enzymes have a preference for cleaving substrates with N-terminal amino acid residues with small non-polar R-groups, as expected for M1 type APs. Furthermore, the two APs were inhibited by the antibiotic puromycin which is used to distinguish membrane bound APs from cytosolic M1 type APs, both of which have similar substrate specificities (Gros *et al.*, 1985). Like the APs of the parasites *F. hepatica*, *L. amazonensis*, *P. falciparum* and *T. gondii* (Morty and Morehead, 2002; Acosta et al., 2007; Stack et al., 2007; Jia et al., 2010), AP1 and AP2 of *T. congolense* proved to be cytosolic as confirmed by the close to

neutral pH optima for activity (pH 7.5 and 8.0 respectively) and by immunofluorescence microscopy.

In order to study the role of these two APs in *T. congolense*, RNAi was conducted. Firstly, it was necessary to identify whether the APs were differentially regulated during the lifecycle of the trypanosome. Western blotting, activity assays using H-Ala-AMC as a substrate and immunofluorescence microscopy of procyclic, metacyclic, epimastigote and bloodstream form parasites showed that AP1 and AP2 were found across all life cycle stages of *T. congolense*. Subsequently, cultures of procyclic trypanosomes were transfected with vectors allowing tetracycline inducible down-regulation of either AP1, AP2 or both AP1 and AP2 expression. No changes in trypanosome morphology or growth were observed in any of the RNAi induced strains. The strain in which AP1 and AP2 were down-regulated did show a slight decrease in growth rate when compared to the wild type culture, but no further differences were observed. Although it is possible to induce differentiation in *T. congolense* cultures *in vivo* and obtain the different life cycle forms, it proved to be impossible with the AP RNAi induced strains. Production of epimastigotes (and subsequently metacyclics) from procyclics of the AP RNAi strains was performed without incident. However, bloodstream form parasites of the AP RNAi strains did not survive in culture despite repeated attempts. The survival of similar strains of *T. congolense* used as positive controls suggest that this inability to survive as cultured bloodstream forms could be an interesting phenotype of the down-regulation of the APs. From this result it can be hypothesised that the presence of the APs is not necessary for the survival of the insect forms of the parasite, but may be essential for survival and growth of the bloodstream form. If this is indeed the case, these APs could prove to be useful targets for future drug development.

Immunisation of mice with recombinant refolded AP1 and AP2 was performed and the antibody response to immunisation measured by ELISA. Although the immunised mice did make specific antibodies against the APs, infection with *T. congolense* (strain IL1180) resulted in high levels of parasitaemia two days after infection and subsequent death of the immunised mice four to five days after infection i.e. no different to the control mice not immunised with APs. Since there are other pathogenic factors of *T. congolense*, including the various cysteine peptidases identified by this study, it is not entirely surprising that immunisation with AP1 and AP2 did not provide protection from infection with *T. congolense* (strain IL1180).

The role of these variant CPs in trypanosomes and trypanosomosis needs to be elucidated and the function of the individual congopain-like peptidases established. Further studies, possibly using gene disruption technology (Mottram *et al.*, 1996) need to be performed to determine the precise role of the multiple CPs *in vivo.* More work needs to be conducted concerning the level of expression of the APs in bloodstream form parasites, as well as establishing the role of these enzymes in extra-cellular parasites possibly with the use of gene knock-out studies, rather than RNAi. Furthermore, it might be feasible to try an alternate expression system, such as baculovirus, for the production of the recombinant aminopeptidases, thus avoiding the high losses of protein which occur during refolding. Once large scale production and purification of the active APs is possible, it would be interesting to perform a vaccination experiment using both the cysteine peptidases (congopain and at least one variant) together with the two APs. Additionally, using a more suitable animal for the vaccination trial, such as goats or even cattle, could give a more realistic indication of the protective potential of these two groups of proteins. Lastly, more immune-proteomic methods need to be investigated to identify all pathogenic factors of *T. congolense.*

The present study has established that the cathepsin L-like cysteine peptidases and two M1 type aminopeptidases are potential pathogenic factors of *T. congolense.* Furthermore, multiple variants of congopain genes are present in the genome, and the recombinant expression of these genes results in peptidases with slight differences in substrate specificities, active at physiological pH. It is apparent that immunisation with either the CPs or the APs alone does not provide protection against challenge with *T. congolense* (strain IL1180) in mice. Furthermore, in terms of an anti-disease vaccine, it is unlikely that the differences in primary sequence of the variant CPs result in a significant difference of epitopes compared to congopain. Therefore, it is improbable that the lack of total protection from trypanosomosis produced by immunisation with CP1 and congopain (Authié *et al.*, 2001) is a result of the presence of the multiple congopain-like CPs analysed in the present study. It is far more likely that there are other pathogenic factors that can be identified by the immuno-proteomic methods outlined here, which need to be included in a multi-component anti-disease vaccine to provide complete protection against AAT.

REFERENCES

Abdulla, M. H., O'Brien, T., Mackey, Z. B., Sajid, M., Grab, D. J. and McKerrow, J. H. 2008. RNA interference of *Trypanosoma brucei* cathepsin B and L affects disease progression in a mouse model. PLoS Negl Trop Dis 2, e298.

Acosta, D., Cancela, M., Piacenza, L., Roche, L., Carmona, C. and Tort, J. F. 2008. *Fasciola hepatica* leucine aminopeptidase, a promising candidate for vaccination against ruminant fasciolosis. Mol Biochem Parasitol 158, 52-64.

Adams, E. R., Hamilton, P. B. and Gibson, W. C. 2010. African trypanosomes: celebrating diversity. Trends Parasitol 26, 324-328.

Aksoy, S. 2003. Control of tsetse flies and trypanosomes using molecular genetics. Vet Parasitol 115, 125-145.

Alcala-Canto, Y., Ibarra-Velarde, F., Sumano-Lopez, H., Gracia-Mora, J. and Alberti-Navarro, A. 2007. Effect of a cysteine protease inhibitor on *Fasciola hepatica* (liver fluke) fecundity, egg viability, parasite burden, and size in experimentally infected sheep. Parasitol Res 100, 461-465.

Allary, M., Schrével, J. and Florent, I. 2002. Properties, stage-dependent expression and localization of *Plasmodium falciparum* M1 family zinc-aminopeptidase. Parasitology 125, 1-10.

Aloulou, A., Grandval, P., De Caro, J., De Caro, A. and Carriere, F. 2006. Constitutive expression of human pancreatic lipase-related protein 1 in *Pichia pastoris*. Protein Expr Purif 47, 415-421.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. 1990. Basic local alignment search tool. J Mol Biol 215, 403-410.

Anene, B. M., Onah, D. N. and Nawa, Y. 2001. Drug resistance in pathogenic African trypanosomes: what hopes for the future? Vet Parasitol 96, 83-100.

Antoine-Moussiaux, N., Büscher, P. and Desmecht, D. 2009. Host-parasite interactions in trypanosomiasis: on the way to an antidisease strategy. Infect Immun 77, 1276-1284.

Antoine-Moussiaux, N., Magez, S. and Desmecht, D. 2008. Contributions of experimental mouse models to the understanding of African trypanosomiasis. Trends Parasitol 24, 411-418.

Aoki, S., Li, J., Itagaki, S., Okech, B. A., Egwang, T. G., Matsuoka, H., Palacpac, N. M., Mitamura, T. and Horii, T. 2002. Serine repeat antigen (SERA5) is predominantly expressed among the SERA multigene family of *Plasmodium falciparum*, and the acquired antibody titers correlate with serum inhibition of the parasite growth. J Biol Chem 277, 47533-47540.

Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195-201.

Atkinson, H. J., Babbitt, P. C. and Sajid, M. 2009. The global cysteine peptidase landscape in parasites. Trends Parasitol 25, 573-581.

Atyame Nten, C. M., Sommerer, N., Rofidal, V., Hirtz, C., Rossignol, M., Cuny, G., Peltier, J. B. and Geiger, A. 2010. Excreted/secreted proteins from trypanosome procyclic strains. J Biomed Biotechnol 2010, 212-220.

Authié, E. 1994. Trypanosomiasis and trypanotolerance in cattle: a role for congopain? Parasitol Today 10, 360-364.

Authié, E., Boulangé, A., Muteti, D., Lalmanach, G., Gauthier, F. and Musoke, A. J. 2001. Immunisation of cattle with cysteine proteinases of *Trypanosoma congolense*: targetting the disease rather than the parasite. Int J Parasitol 31, 1429-1433.

Authié, E., Duvallet, G., Robertson, C. and Williams, D. J. 1993a. Antibody responses to a 33 kDa cysteine protease of *Trypanosoma congolense*: relationship to 'trypanotolerance' in cattle. Parasite Immunol 15, 465-474.

Authié, E., Muteti, D. K., Mbawa, Z. R., Lonsdale-Eccles, J. D., Webster, P. and Wells, C. W. 1992. Identification of a 33-kilodalton immunodominant antigen of *Trypanosoma congolense* as a cysteine protease. Mol Biochem Parasitol 56, 103-116.

Authié, E., Muteti, D. K. and Williams, D. J. 1993b. Antibody responses to invariant antigens of *Trypanosoma congolense* in cattle of differing susceptibility to trypanosomiasis. Parasite Immunol 15, 101-111.

Bannai, H., Sakurai, T., Inoue, N., Sugimoto, C. and Igarashi, I. 2003. Cloning and Expression of mitochondrial heat shock protein 70 of *Trypanosoma congolense* and potential use as a diagnostic antigen. Clin Diagn Lab Immunol 10, 926-933.

Baral, T. N. 2010. Immunobiology of African trypanosomes: need of alternative interventions. J Biomed Biotechnol 2010, 389153.

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

Barrett, A. J. and Kirschke, H. 1981. Cathepsin B, Cathepsin H, and cathepsin L. Methods Enzymol 80 Pt C, 535-561.

Barrett, A. J. and Rawlings, N. D. 2001. Evolutionary lines of cysteine peptidases. Biol Chem 382, 727-733.

Barrett, A. J. and Rawlings, N. D. 2004a. Classification of aminopeptidases. In: Barrett, A. J., Rawlings, N. D., Woessner, J. F. (Eds.), Handbook of proteolytic enzymes. Elsevier Academic Press, London, pp. 231-235.

Barrett, A. J. and Rawlings, N. D. 2004b. Introduction: The clans and families of cysteine peptidases. In: Barrett, A. J., Rawlings, N. D., Woessner, J. F. (Eds.), Handbook of proteolytic enzymes. Elsevier, London, pp. 1051-1071.

Barry, D. and Carrington, M. 2004. Antigenic Variation. In: Maudlin, I., Holmes, P. H., Miles, M. A. (Eds.), The Trypanosomiases. CAB International, Wallingford, pp. 25-37.

Barry, J. D., Marcello, L., Morrison, L. J., Read, A. F., Lythgoe, K., Jones, N., Carrington, M., Blandin, G., Bohme, U., Caler, E., Hertz-Fowler, C., Renauld, H., El-Sayed, N. and Berriman, M. 2005. What the genome sequence is revealing about trypanosome antigenic variation. Biochem Soc Trans 33, 986-989.

Bastin, P., Ellis, K., Kohl, L. and Gull, K. 2000a. Flagellum ontogeny in trypanosomes studied via an inherited and regulated RNA interference system. J Cell Sci 113 3321-3328.

Bastin, P., Pullen, T. J., Moreira-Leite, F. F. and Gull, K. 2000b. Inside and outside of the trypanosome flagellum:a multifunctional organelle. Microbes Infect 2, 1865-1874.

Beckham, S. A., Piedrafita, D., Phillips, C. I., Samarawickrema, N., Law, R. H., Smooker, P. M., Quinsey, N. S., Irving, J. A., Greenwood, D., Verhelst, S. H., Bogyo, M., Turk, B., Coetzer, T. H., Wijeyewickrema, L. C., Spithill, T. W. and Pike, R. N. 2009. A major cathepsin B protease from the liver fluke *Fasciola hepatica* has atypical active site features and a potential role in the digestive tract of newly excysted juvenile parasites. Int J Biochem Cell Biol 41, 1601-1612.

Bellofatto, V. and Palenchar, J. B. 2008. RNA interference as a genetic tool in trypanosomes. Methods Mol Biol 442, 83-94.

Bengaly, Z., Sidibe, I., Ganaba, R., Desquesnes, M., Boly, H. and Sawadogo, L. 2002. Comparative pathogenicity of three genetically distinct types of *Trypanosoma congolense* in cattle: clinical observations and haematological changes. Vet Parasitol 108, 1-19.

Berthier, D., Quere, R., Thevenon, S., Belemsaga, D., Piquemal, D., Marti, J. and Maillard, J. C. 2003. Serial analysis of gene expression (SAGE) in bovine trypanotolerance: preliminary results. Genet Sel Evol 35 Suppl 1, S35-47.

Billington, C. J., Mason, P., Magny, M. C. and Mort, J. S. 2000. The slow-binding inhibition of cathepsin K by its propeptide. Biochem Biophys Res Commun 276, 924-929.

Bizaaré, L. C. 2008. Evaluation of congopain and oligopeptidase B as anti-disease vaccines for African trypanosomiasis. School of Biochemistry, Genetics and Microbiology. University of KwaZulu-Natal, Pietermaritzburg, pp. 105.

Black, S. J., Sicard, E. L., Murphy, N. and Nolan, D. 2001. Innate and acquired control of trypanosome parasitaemia in Cape buffalo. Int J Parasitol 31, 562-565.

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

Bossard, G., Boulangé, A., Holzmuller, P., Thevenon, S., Patrel, D. and Authié, E. 2010. Serodiagnosis of bovine trypanosomosis based on HSP70/BiP inhibition ELISA. Vet Parasitol 173, 39-47.

Botbol, V. and Scornik, O. A. 1991. Measurement of instant rates of protein degradation in the livers of intact mice by the accumulation of bestatin-induced peptides. J Biol Chem 266, 2151-2157.

Boulangé, A., Katende, J. and Authié, E. 2002. *Trypanosoma congolense*: expression of a heat shock protein 70 and initial evaluation as a diagnostic antigen for bovine trypanosomosis. Exp Parasitol 100, 6-11.

Boulangé, A., Serveau, C., Brillard, M., Minet, C., Gauthier, F., Diallo, A., Lalmanach, G. and Authié, E. 2001. Functional expression of the catalytic domains of two cysteine proteinases from *Trypanosoma congolense*. Int J Parasitol 31, 1435-1440.

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

Bouyer, J., Stachurski, F., Gouro, A. S. and Lancelot, R. 2009. Control of bovine trypanosomosis by restricted application of insecticides to cattle using footbaths. Vet Parasitol 161, 187-193.

Brak, K., Kerr, I. D., Barrett, K. T., Fuchi, N., Debnath, M., Ang, K., Engel, J. C., McKerrow, J. H., Doyle, P. S., Brinen, L. S. and Ellman, J. A. 2010. Nonpeptidic tetrafluorophenoxymethyl ketone cruzain inhibitors as promising new leads for Chagas disease chemotherapy. J Med Chem 53, 1763-1773.

Brenchley, R., Tariq, H., McElhinney, H., Szoor, B., Huxley-Jones, J., Stevens, R., Matthews, K. and Tabernero, L. 2007. The TriTryp Phosphatome: analysis of the protein phosphatase catalytic domains. BMC Genomics 8, 434.

Briand, J. P., Muller, S. and Van Regenmortel, M. H. 1985. Synthetic peptides as antigens: pitfalls of conjugation methods. J Immunol Methods 78, 59-69.

Broadhead, R., Dawe, H. R., Farr, H., Griffiths, S., Hart, S. R., Portman, N., Shaw, M. K., Ginger, M. L., Gaskell, S. J., McKean, P. G. and Gull, K. 2006. Flagellar motility is required for the viability of the bloodstream trypanosome. Nature 440, 224-227.

Brown, K. 2008. From Ubombo to Mkhuzi: disease, colonial science, and the control of Nagana (livestock trypanosomosis) in Zululand, South Africa, C. 1894-1953. J Hist Med Allied Sci 63, 285-322.

Bruce, D. 1914. Classification of the African Trypanosomes pathogenic to man and domestic animals. Trans R Soc Trop Med Hyg 8, 1-22.

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

Bryson, K., Besteiro, S., McGachy, H. A., Coombs, G. H., Mottram, J. C. and Alexander, J. 2009. Overexpression of the natural inhibitor of cysteine peptidases in *Leishmania mexicana* leads to reduced virulence and a Th1 response. Infect Immun 77, 2971-2978.

Bulaj, G. 2005. Formation of disulfide bonds in proteins and peptides. Biotechnol Adv 23, 87-92.

Burgess, R. R. 2009. Refolding solubilized inclusion body proteins. Methods Enzymol 463, 259-282.

Cabrera, L., De Witte, J., Victor, B., Vermeiren, L., Zimic, M., Brandt, J. and Geysen, D. 2009. Specific detection and identification of African trypanosomes in bovine peripheral blood by means of a PCR-ELISA assay. Vet Parasitol 164, 111-117.

Cabrita, L. D. and Bottomley, S. P. 2004. Protein expression and refolding--a practical guide to getting the most out of inclusion bodies. Biotechnol Annu Rev 10, 31-50.

Caffrey, C. R., Hansell, E., Lucas, K. D., Brinen, L. S., Alvarez Hernandez, A., Cheng, J., Gwaltney, S. L., 2nd, Roush, W. R., Stierhof, Y. D., Bogyo, M., Steverding, D. and McKerrow, J. H. 2001. Active site mapping, biochemical properties and subcellular localization of rhodesain, the major cysteine protease of *Trypanosoma brucei rhodesiense*. Mol Biochem Parasitol 118, 61-73.

Caffrey, C. R. and Steverding, D. 2009. Kinetoplastid papain-like cysteine peptidases. Mol Biochem Parasitol 167, 12-19.

Cancela, M., Acosta, D., Rinaldi, G., Silva, E., Duran, R., Roche, L., Zaha, A., Carmona, C. and Tort, J. F. 2008. A distinctive repertoire of cathepsins is expressed by juvenile invasive *Fasciola hepatica*. Biochimie 90, 1461-1475.

Cármenes, R. S., Freije, J. P., Molina, M. M. and Martín, J. M. 1989. Predict7, a program for protein structure prediction. Biochem Biophys Res Commun 159, 687-693.

Carmona, E., Dufour, E., Plouffe, C., Takebe, S., Mason, P., Mort, J. S. and Ménard, R. 1996. Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases. Biochemistry 35, 8149-8157.

Cazorla, S. I., Becker, P. D., Frank, F. M., Ebensen, T., Sartori, M. J., Corral, R. S., Malchiodi, E. L. and Guzmán, C. A. 2008. Oral vaccination with *Salmonella enterica* as a cruzipain-DNA delivery system confers protective immunity against *Trypanosoma cruzi*. Infect Immun 76, 324-333.

Cazorla, S. I., Frank, F. M. and Malchiodi, E. L. 2009. Vaccination approaches against *Trypanosoma cruzi* infection. Expert Rev Vaccines 8, 921-935.

Cazzulo, J., Stoka, V. and Turk, V. 2001. The major cysteine proteinase of *Trypanosoma cruzi*: a valid target for chemotherapy of Chagas disease. Curr Pharm Des 7, 1143-1156.

Cereghino, G. P., Cereghino, J. L., Ilgen, C. and Cregg, J. M. 2002. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. Curr Opin Biotechnol 13, 329-332.

Chagas, J. R., Authié, E., Serveau, C., Lalmanach, G., Juliano, L. and Gauthier, F. 1997. A comparison of the enzymatic properties of the major cysteine proteinases from *Trypanosoma congolense* and *Trypanosoma cruzi*. Mol Biochem Parasitol 88, 85-94.

Chandler, F. W. and Watts, J. C. 1988. Immunofluorescence as an adjunct to the histopathologic diagnosis of Chagas' disease. J Clin Microbiol 26, 567-569.

Chen, X., Chong, C. R., Shi, L., Yoshimoto, T., Sullivan, D. J., Jr. and Liu, J. O. 2006. Inhibitors of *Plasmodium falciparum* methionine aminopeptidase 1b possess antimalarial activity. Proc Natl Acad Sci U S A 103, 14548-14553.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 31, 3497-3500.

Claes, F., Büscher, P., Touratier, L. and Goddeeris, B. M. 2005. *Trypanosoma equiperdum*: master of disguise or historical mistake? Trends Parasitol 21, 316-321.

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

Cohen, P. 1989. The structure and regulation of protein phosphatases. Annu Rev Biochem 58, 453-508.

Correnti, J. M., Brindley, P. J. and Pearce, E. J. 2005. Long-term suppression of cathepsin B levels by RNA interference retards schistosome growth. Mol Biochem Parasitol 143, 209-215.

Courtin, D., Berthier, D., Thevenon, S., Dayo, G. K., Garcia, A. and Bucheton, B. 2008. Host genetics in African trypanosomiasis. Infect Genet Evol 8, 229-238.

Coustou, V., Guegan, F., Plazolles, N. and Baltz, T. 2010. Complete in vitro life cycle of *Trypanosoma congolense*: development of genetic tools. PLoS Negl Trop Dis 4, e618.

Cuervo, P., Domont, G. B. and De Jesus, J. B. 2010. Proteomics of trypanosomatids of human medical importance. J Proteomics 73, 845-867.

Cygler, M. and Mort, J. S. 1997. Proregion structure of members of the papain superfamily. Mode of inhibition of enzymatic activity. Biochimie 79, 645-652.

d'Ieteren, G. D., Authié, E., Wissocq, N. and Murray, M. 1998. Trypanotolerance, an option for sustainable livestock production in areas at risk from trypanosomosis. Rev Sci Tech 17, 154-175.

d'Ieteren, G. D. and Kimani, K. 2007. Indigenous genetic resources: A sustainable and environmentally friendly option for livestock production in areas at risk from trypanosomes. Science in Africa.

Dalton, J. P. and Mulcahy, G. 2001. Parasite vaccines - a reality? Vet Parasitol 98, 149-167.

DeLano, W. L. 2004. PyMOL. The PyMOL molecular graphics system. DeLano Scientific, San Carlos.

Delespaux, V., Geysen, D., Van den Bossche, P. and Geerts, S. 2008. Molecular tools for the rapid detection of drug resistant in animal trypanosomes. Trends in Parasitology 24, 236-242.

Delespaux, V., Vitouley, H. S., Marcotty, T., Speybroeck, N., Berkvens, D., Roy, K., Geerts, S. and Van den Bossche, P. 2010. Chemosensitization of *Trypanosoma congolense* strains resistant to isometamidium chloride by tetracyclines and enrofloxacin. PLoS Negl Trop Dis 4, e828.

Denise, H., McNeil, K., Brooks, D. R., Alexander, J., Coombs, G. H. and Mottram, J. C. 2003. Expression of multiple CPB genes encoding cysteine proteases is required for *Leishmania mexicana* virulence in vivo. Infect Immun 71, 3190-3195.

Derbyshire, J. B. and Nielsen, N. O. 1997. Edward Watson and the eradication of dourine in Canada. Can Vet J 38, 582-586.

Desquesnes, M. and Dávila, A. M. 2002. Applications of PCR-based tools for detection and identification of animal trypanosomes: a review and perspectives. Vet Parasitol 109, 213-231.

Dixit, A. K., Dixit, P. and Sharma, R. L. 2008. Immunodiagnostic/protective role of cathepsin L cysteine proteinases secreted by *Fasciola* species. Vet Parasitol 154, 177-184.

Doerig, C., Abdi, A., Bland, N., Eschenlauer, S., Dorin-Semblat, D., Fennell, C., Halbert, J., Holland, Z., Nivez, M. P., Semblat, J. P., Sicard, A. and Reininger, L. 2010. Malaria: targeting parasite and host cell kinomes. Biochim Biophys Acta 1804, 604-612.

Donelson, J. E. 2003. Antigenic variation and the African trypanosome genome. Acta Trop 85, 391-404.

Downey, N. and Donelson, J. E. 1999. Expression of foreign proteins in *Trypanosoma congolense.* Mol Biochem Parasitol 104, 39-53.

Duggan, A. J. 1977. Bruce and the African Trypanosomes. Am J Trop Med Hyg 26, 1080-1083.

Duschak, V. G. and Couto, A. S. 2009. Cruzipain, the major cysteine protease of *Trypanosoma cruzi:* a sulfated glycoprotein antigen as relevant candidate for vaccine development and drug target. A review. Curr Med Chem 16, 3174-3202.

Dvořák, J., Delcroix, M., Rossi, A., Vopálenský, V., Pospíšek, M., Šedinová, M., Mikeš, L., Sajid, M., Sali, A., McKerrow, J. H., Horák, P. and Caffrey, C. R. 2005. Multiple cathepsin B isoforms in schistosomula of *Trichobilharzia regenti*: identification, characterisation and putative role in migration and nutrition. Int J Parasitol 35, 895-910.

Eakin, A. E., McGrath, M. E., McKerrow, J. H., Fletterick, R. J. and Craik, C. S. 1993. Production of crystallizable cruzain, the major cysteine protease from *Trypanosoma cruzi*. J Biol Chem 268, 6115-6118.

Eakin, A. E., Mills, A. A., Harth, G., McKerrow, J. H. and Craik, C. S. 1992. The sequence, organization, and expression of the major cysteine protease (cruzain) from *Trypanosoma cruzi*. J Biol Chem 267, 7411-7420.

Eisler, M. C., Dwinger, R. H., Majiwa, P. A. and Picozzi, K. 2004. Diagnosis and Epidemiology of African Animal Trypanosomiasis. In: Maudlin, I., Holmes, P. H., Miles, M. A. (Eds.), The Trypanosomiases. CAB International, Wallingford, pp. 253-267.

El-Sayed, N. M., Hegde, P., Quackenbush, J., Melville, S. E. and Donelson, J. E. 2000. The African trypanosome genome. Int J Parasitol 30, 329-345.

el Kouni, M. H. 2003. Potential chemotherapeutic targets in the purine metabolism of parasites. Pharmacol Ther 99, 283-309.

Ellis, K. J. and Morrison, J. F. 1982. Buffers of constant ionic strength for studying pH-dependent processes. Methods Enzymol 87, 405-426.

Engel, J. C., Doyle, P. S., Palmer, J., Hsieh, I., Bainton, D. F. and McKerrow, J. H. 1998. Cysteine protease inhibitors alter Golgi complex ultrastructure and function in *Trypanosoma cruzi*. J Cell Sci 111 (Pt 5), 597-606.

Englund, P. T., Hajduk, S. L. and Marini, J. C. 1982. The molecular biology of trypanosomes. Annu Rev Biochem 51, 695-726.

Engstler, M., Schauer, R. and Brun, R. 1995. Distribution of developmentally regulated trans-sialidases in the Kinetoplastida and characterization of a shed trans-sialidase activity from procyclic *Trypanosoma congolense*. Acta Trop 59, 117-129.

Enyaru, J. C., Ouma, J. O., Malele, II, Matovu, E. and Masiga, D. K. 2010. Landmarks in the evolution of technologies for identifying trypanosomes in tsetse flies. Trends Parasitol 26, 388-394.

Ettari, R., Bova, F., Zappala, M., Grasso, S. and Micale, N. 2010. Falcipain-2 inhibitors. Med Res Rev 30, 136-167.

Field, M. C. and Carrington, M. 2009. The trypanosome flagellar pocket. Nat Rev Microbiol 7, 775-786.

Finelle, P. 1979. African animal trypanosomiasis Part 1: Disease and Chemotherapy. FAO. ILRAD.

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

Fish, W. R., Nkhungulu, Z. M., Muriuki, C. W., Ndegwa, D. M., Lonsdale-Eccles, J. D. and Steyaert, J. 1995. Primary structure and partial characterization of a life-cycle-regulated cysteine protease from *Trypanosoma (Nannomonas) congolense*. Gene 161, 125-128.

Flipo, M., Beghyn, T., Leroux, V., Florent, I., Deprez, B. P. and Deprez-Poulain, R. F. 2007. Novel selective inhibitors of the zinc plasmodial aminopeptidase PfA-M1 as potential antimalarial agents. J Med Chem 50, 1322-1334.

Florent, I., Derhy, Z., Allary, M., Monsigny, M., Mayer, R. and Schrével, J. 1998. A *Plasmodium falciparum* aminopeptidase gene belonging to the M1 family of zinc-metallopeptidases is expressed in erythrocytic stages. Mol Biochem Parasitol 97, 149-160.

Gardiner, D. L., Skinner-Adams, T. S., Brown, C. L., Andrews, K. T., Stack, C. M., McCarthy, J. S., Dalton, J. P. and Trenholme, K. R. 2009. *Plasmodium falciparum*: new molecular targets with potential for antimalarial drug development. Expert Rev Anti Infect Ther 7, 1087-1098.

Gardiner, D. L., Trenholme, K. R., Skinner-Adams, T. S., Stack, C. M. and Dalton, J. P. 2006. Overexpression of leucyl aminopeptidase in *Plasmodium falciparum* parasites. Target for the antimalarial activity of bestatin. J Biol Chem 281, 1741-1745.

Geerts, S. and Gryseels, B. 2000. Drug resistance in human helminths: current situation and lessons from livestock. Clin Microbiol Rev 13, 207-222.

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

Geigy, R. and Kauffmann, M. 1973. Sleeping sickness survey in the Serengeti area (Tanzania) 1971. I. Examination of large mammals for trypanosomes. Acta Trop 30, 12-23.

Georgiadou, D., Hearn, A., Evnouchidou, I., Chroni, A., Leondiadis, L., York, I. A., Rock, K. L. and Stratikos, E. 2010. Placental leucine aminopeptidase efficiently generates mature antigenic peptides in vitro but in patterns distinct from endoplasmic reticulum aminopeptidase 1. J Immunol 185, 1584-1592.

Gibson, W. 2007. Resolution of the species problem in African trypanosomes. Int J Parasitol 37, 829-838.

Gibson, W., Peacock, L., Ferris, V., Williams, K. and Bailey, M. 2008. The use of yellow fluorescent hybrids to indicate mating in *Trypanosoma brucei*. Parasit Vectors 1, 4.

Godat, E., Chowdhury, S., Lecaille, F., Belghazi, M., Purisima, E. O. and Lalmanach, G. 2005. Inhibition of a cathepsin L-like cysteine protease by a chimeric propeptide-derived inhibitor. Biochemistry 44, 10486-10493.

Goldring, J. P., Thobakgale, C., Hiltunen, T. and Coetzer, T. H. 2005. Raising antibodies in chickens against primaquine, pyrimethamine, dapsone, tetracycline, and doxycycline. Immunol Invest 34, 101-114.

Gomez, C., Esther Ramirez, M., Calixto-Galvez, M., Medel, O. and Rodriguez, M. A. 2010. Regulation of gene expression in protozoa parasites. J Biomed Biotechnol 2010, 726045.

Gooding, R. H. and Krafsur, E. S. 2005. Tsetse genetics: contributions to biology, systematics, and control of tsetse flies. Annu Rev Entomol 50, 101-123.

Grams, R., Vichasri-Grams, S., Sobhon, P., Upatham, E. S. and Viyanant, V. 2001. Molecular cloning and characterization of cathepsin L encoding genes from *Fasciola gigantica*. Parasitol Int 50, 105-114.

Grébaut, P., Chuchana, P., Brizard, J. P., Demettre, E., Seveno, M., Bossard, G., Jouin, P., Vincendeau, P., Bengaly, Z., Boulangé, A., Cuny, G. and Holzmuller, P. 2009. Identification of total and differentially expressed excreted-secreted proteins from *Trypanosoma congolense* strains exhibiting different virulence and pathogenicity. Int J Parasitol.

Gros, C., Giros, B. and Schwartz, J. C. 1985. Purification of membrane-bound aminopeptidase from rat brain: identification of aminopeptidase M. Neuropeptides 5, 485-488.

Guay, J., Falgueyret, J. P., Ducret, A., Percival, M. D. and Mancini, J. A. 2000. Potency and selectivity of inhibition of cathepsin K, L and S by their respective propeptides. Eur J Biochem 267, 6311-6318.

Guirnalda, P., Murphy, N. B., Nolan, D. and Black, S. J. 2007. Anti-*Trypanosoma brucei* activity in Cape buffalo serum during the cryptic phase of parasitemia is mediated by antibodies. Int J Parasitol 37, 1391-1399.

Haag, J., O'HUigin, C. and Overath, P. 1998. The molecular phylogeny of trypanosomes: evidence for an early divergence of the Salivaria. Mol Biochem Parasitol 91, 37-49.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. . Nucl Acids Symp Ser 41, 95-98.

Hambrey, P. N., Mellors, A. and Tizard, I. R. 1981. The phospholipases of pathogenic and non-pathogenic *Trypanosoma* species. Mol Biochem Parasitol 2, 177-186.

Hanotte, O., Tawah, C. L., Bradley, D. G., Okomo, M., Verjee, Y., Ochieng, J. and Rege, J. E. 2000. Geographic distribution and frequency of a taurine *Bos taurus* and an indicine *Bos indicus* Y specific allele amongst sub-saharan African cattle breeds. Mol Ecol 9, 387-396.

Hanquier, J., Sorlet, Y., Desplancq, D., Baroche, L., Ebtinger, M., Lefevre, J. F., Pattus, F., Hershberger, C. L. and Vertes, A. A. 2003. A single mutation in the activation site of

bovine trypsinogen enhances its accumulation in the fermentation broth of the yeast Pichia pastoris. Appl Environ Microbiol 69, 1108-1113.

Hatta, T., Kazama, K., Miyoshi, T., Umemiya, R., Liao, M., Inoue, N., Xuan, X., Tsuji, N. and Fujisaki, K. 2006. Identification and characterisation of a leucine aminopeptidase from the hard tick *Haemaphysalis longicornis*. Int J Parasitol 36, 1123-1132.

Hatta, T., Umemiya, R., Liao, M., Gong, H., Harnnoi, T., Tanaka, M., Miyoshi, T., Boldbaatar, D., Battsetseg, B., Zhou, J., Xuan, X., Tsuji, N., Taylor, D. and Fujisaki, K. 2007. RNA interference of cytosolic leucine aminopeptidase reduces fecundity in the hard tick, *Haemaphysalis longicornis*. Parasitol Res 100, 847-854.

Hattori, A. and Tsujimoto, M. 2004. Processing of antigenic peptides by aminopeptidases. Biol Pharm Bull 27, 777-780.

Heby, O., Persson, L. and Rentala, M. 2007. Targeting the polyamine biosynthetic enzymes: a promising approach to therapy of African sleeping sickness, Chagas' disease, and leishmaniasis. Amino Acids 33, 359-366.

Heussen, C. J. and Dowdle, E. B. 1980. Electrophoretic analysis of plasminogen activators in polyacrylamide gels containing sodium dodecyl sulfate and copolymerised substrates. Anal Biochem 102, 196-202.

Hide, G. 2008. Visualizing trypanosome sex. Trends Parasitol 24, 425-428.

Hide, M. and Banuls, A. L. 2008. Polymorphisms of *cpb* multicopy genes in the *Leishmania (Leishmania) donovani* complex. Trans R Soc Trop Med Hyg 102, 105-106.

Hodder, A. N., Drew, D. R., Epa, V. C., Delorenzi, M., Bourgon, R., Miller, S. K., Moritz, R. L., Frecklington, D. F., Simpson, R. J., Speed, T. P., Pike, R. N. and Crabb, B. S. 2003. Enzymic, phylogenetic, and structural characterization of the unusual papain-like protease domain of *Plasmodium falciparum* SERA5. J Biol Chem 278, 48169-48177.

Holmes, P. H., Eisler, M. C. and Geerts, S. 2004. Current Chemotherapy of Animal Trypanosomiasis. In: Maudlin, I., Holmes, P. H., Miles, M. A. (Eds.), The Trypanosomiases. CAb International, Wallingford, pp. 431-444.

Holzmuller, P., Biron, D. G., Courtois, P., Koffi, M., Bras-Goncalves, R., Daulouede, S., Solano, P., Cuny, G., Vincendeau, P. and Jamonneau, V. 2008a. Virulence and pathogenicity patterns of *Trypanosoma brucei gambiense* field isolates in experimentally infected mouse: differences in host immune response modulation by secretome and proteomics. Microbes Infect 10, 79-86.

Holzmuller, P., Grébaut, P., Peltier, J. B., Brizard, J. P., Perrone, T., Gonzatti, M., Bengaly, Z., Rossignol, M., Aso, P. M., Vincendeau, P., Cuny, G., Boulangé, A. and Frutos, R. 2008b. Secretome of animal trypanosomes. Ann N Y Acad Sci 1149, 337-342.

Holzmuller, P., Herder, S., Cuny, G. and De Meeus, T. 2010. From clonal to sexual: a step in *T. congolense* evolution? Trends Parasitol 26, 56-60.

Hooper, N. M. 1994. Families of zinc metalloproteases. FEBS Lett 354, 1-6.

Hotz, H. R., Hartmann, C., Huober, K., Hug, M. and Clayton, C. 1997. Mechanisms of developmental regulation in *Trypanosoma brucei*: a polypyrimidine tract in the 3'-untranslated region of a surface protein mRNA affects RNA abundance and translation. Nucleic Acids Res 25, 3017-3026.

Huson, L. E., Authié, E., Boulangé, A. F., Goldring, J. P. and Coetzer, T. H. 2009. Modulation of the immunogenicity of the *Trypanosoma congolense* cysteine protease, congopain, through complexation with alpha(2)-macroglobulin. Vet Res 40, 52-64.

Inoue, N., Otsu, K., Ferraro, D. M. and Donelson, J. E. 2002. Tetracycline-regulated RNA interference in *Trypanosoma congolense*. Mol Biochem Parasitol 120, 309-313.

Ito, K., Nakajima, Y., Onohara, Y., Takeo, M., Nakashima, K., Matsubara, F., Ito, T. and Yoshimoto, T. 2006. Crystal structure of aminopeptidase N (proteobacteria alanyl aminopeptidase) from *Escherichia coli* and conformational change of methionine 260 involved in substrate recognition. J Biol Chem 281, 33664-33676.

Jayaraj, R., Piedrafita, D., Dynon, K., Grams, R., Spithill, T. W. and Smooker, P. M. 2009. Vaccination against fasciolosis by a multivalent vaccine of stage-specific antigens. Vet Parasitol 160, 230-236.

Jaye, A. B., Nantulya, V. M., Majiwa, P. A., Urakawa, T., Masake, R. A., Wells, C. W. and Ole-MoiYoi, O. 1994. A *Trypanosoma* (*nannomonas*) *congolense*-specific antigen released into the circulation of infected animals is a thiol protease precursor. EMBL accession L25130.

Jia, H., Nishikawa, Y., Luo, Y., Yamagishi, J., Sugimoto, C. and Xuan, X. 2010. Characterization of a leucine aminopeptidase from *Toxoplasma gondii*. Mol Biochem Parasitol 170, 1-6.

Kakundi, E. M. 2008. Molecular analysis of the congopain gene family. School of Biochemistry, Genetics, Microbiology and Plant Pathology. University of KwaZulu-Natal, Pietermaritzburg.

Kesik, M., Jedlina-Panasiuk, L., Kozak-Cieszczyk, M., Plucienniczak, A. and Wedrychowicz, H. 2007. Enteral vaccination of rats against *Fasciola hepatica* using recombinant cysteine proteinase (cathepsin L1). Vaccine 25, 3619-3628.

Khoshgoo, N., Zahedifard, F., Azizi, H., Taslimi, Y., Alonso, M. J. and Rafati, S. 2008. Cysteine proteinase type III is protective against *Leishmania infantum* infection in BALB/c mice and highly antigenic in visceral leishmaniasis individuals. Vaccine 26, 5822-5829.

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L. and Schwede, T. 2009. The SWISS-MODEL Repository and associated resources. Nucleic Acids Res 37, D387-392.

Kikuchi, S. A., Sodre, C. L., Kalume, D. E., Elias, C. G., Santos, A. L., de Nazare Soeiro, M., Meuser, M., Chapeaurouge, A., Perales, J. and Fernandes, O. 2010. Proteomic analysis of two *Trypanosoma cruzi* zymodeme 3 strains. Exp Parasitol 126, 540-551.

Klemba, M. and Goldberg, D. E. 2002. Biological roles of proteases in parasitic protozoa. Annu Rev Biochem 71, 275-305.

Knowles, G. 1993. The effects of arphamenine-A, an inhibitor of aminopeptidases, on in-vitro growth of *Trypanosoma brucei brucei*. J Antimicrob Chemother 32, 172-174.

Kooy, R. F., Hirumi, H., Moloo, S. K., Nantulya, V. M., Dukes, P., Van der Linden, P. M., Duijndam, W. A., Janse, C. J. and Overdulve, J. P. 1989. Evidence for diploidy in metacyclic forms of African trypanosomes. Proc Natl Acad Sci U S A 86, 5469-5472.

Krafsur, E. S. 2009. Tsetse flies: genetics, evolution, and role as vectors. Infect Genet Evol 9, 124-141.

Kristjanson, P. M., Swallow, B. M., Rowlands, G. J., Kruska, R. L. and Leeuw, P. N. d. 1999. Measuring the costs of African animal trypanosomosis, the potential benefits of control and returns to research. Agricultural Systems, 79-98.

Kumar, A., Kumar, K., Korde, R., Puri, S. K., Malhotra, P. and Singh Chauhan, V. 2007. Falcipain-1, a *Plasmodium falciparum* cysteine protease with vaccine potential. Infect Immun 75, 2026-2034.

LaCount, D. J., Bruse, S., Hill, K. L. and Donelson, J. E. 2000. Double-stranded RNA interference in *Trypanosoma brucei* using head-to-head promoters. Mol Biochem Parasitol 111, 67-76.

LaCount, D. J. and Donelson, J. E. 2001. RNA interference in African trypanosomes. Protist 152, 103-111.

Laderach, D., Cerban, F., Motran, C., Vottero de Cima, E. and Gea, S. 1996. *Trypanosoma cruz*i: the major cysteinyl proteinase (cruzipain) is a relevant immunogen of parasite acidic antigens (FIII). Int J Parasitol 26, 1249-1254.

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

Lalmanach, G., Boulangé, A., Serveau, C., Lecaille, F., Scharfstein, J., Gauthier, F. and Authié, E. 2002. Congopain from *Trypanosoma congolense*: drug target and vaccine candidate. Biol Chem 383, 739-749.

Lalmanach, G., Lecaille, F., Chagas, J. R., Authié, E., Scharfstein, J., Juliano, M. A. and Gauthier, F. 1998. Inhibition of trypanosomal cysteine proteinases by their propeptides. J Biol Chem 273, 25112-25116.

Lalor, T. M., Kjeldgaard, M., Shimamoto, G. T., Strickler, J. E., Konigsberg, W. H. and Richards, F. F. 1984. Trypanosome variant-specific glycoproteins: a polygene protein family with multiple folding patterns? Proc Natl Acad Sci U S A 81, 998-1002.

Lamont, G. S., Fox, J. A. and Cross, G. A. 1987. Glycosyl-sn-1,2-dimyristylphosphatidylinositol is the membrane anchor for *Trypanosoma equiperdum* and *T.* (Nannomonas) *congolense* variant surface glycoproteins. Mol Biochem Parasitol 24, 131-136.

Lange, C. and Rudolph, R. 2009. Suppression of protein aggregation by L-arginine. Curr Pharm Biotechnol 10, 408-414.

Lanham, S. M. and Godfrey, D. G. 1970. Isolation of salivarian trypanosomes from man and other mammals using DEAE-cellulose. Exp Parasitol 28, 521-534.

Laslo, A. C., Ganea, E. and Obinger, C. 2009. Refolding of hexameric porcine leucine aminopeptidase using a cationic detergent and dextrin-10 as artificial chaperones. J Biotechnol 140, 162-168.

Laurière, M. 1993. A semidry electroblotting system efficiently transfers both high- and low-molecular-weight proteins separated by SDS-PAGE. Anal Biochem 212, 206-211.

Laustsen, P. G., Vang, S. and Kristensen, T. 2001. Mutational analysis of the active site of human insulin-regulated aminopeptidase. Eur J Biochem 268, 98-104.

Lecaille, F., Authié, E., Moreau, T., Serveau, C., Gauthier, F. and Lalmanach, G. 2001a. Subsite specificity of trypanosomal cathepsin L-like cysteine proteases. Probing the S2 pocket with phenylalanine-derived amino acids. Eur J Biochem 268, 2733-2741.

Lecaille, F., Cotton, J., McKerrow, J. H., Ferrer-Di Martino, M., Boll-Bataille, E., Gauthier, F. and Lalmanach, G. 2001b. Reversible inhibition of cathepsin L-like proteases by 4-mer pseudopeptides. FEBS Lett 507, 362-366.

Lecaille, F., Kaleta, J. and Brömme, D. 2002. Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. Chem Rev 102, 4459-4488.

Leppert, B. J., Mansfield, J. M. and Paulnock, D. M. 2007. The soluble variant surface glycoprotein of African trypanosomes is recognized by a macrophage scavenger receptor and induces I kappa B alpha degradation independently of TRAF6-mediated TLR signaling. J Immunol 179, 548-556.

Li, S. Q., Fung, M. C., Reid, S. A., Inoue, N. and Lun, Z. R. 2007. Immunization with recombinant beta-tubulin from *Trypanosoma evansi* induced protection against *T. evansi, T. equiperdum* and *T. b. brucei* infection in mice. Parasite Immunol 29, 191-199.

Lilley, C. J., Goodchild, S. A., Atkinson, H. J. and Urwin, P. E. 2005. Cloning and characterisation of a *Heterodera glycines* aminopeptidase cDNA. Int J Parasitol 35, 1577-1585.

Lima, A. P., dos Reis, F. C., Serveau, C., Lalmanach, G., Juliano, L., Ménard, R., Vernet, T., Thomas, D. Y., Storer, A. C. and Scharfstein, J. 2001. Cysteine protease isoforms from *Trypanosoma cruzi*, cruzipain 2 and cruzain, present different substrate preference and susceptibility to inhibitors. Mol Biochem Parasitol 114, 41-52.

Liu, B., Liu, Y., Motyka, S. A., Agbo, E. E. and Englund, P. T. 2005. Fellowship of the rings: the replication of kinetoplast DNA. Trends Parasitol 21, 363-369.

Lonsdale-Eccles, J. D. and Grab, D. J. 2002. Trypanosome hydrolases and the blood-brain barrier. Trends Parasitol 18, 17-19.

Lonsdale-Eccles, J. D., Mpimbaza, G. W., Nkhungulu, Z. R., Olobo, J., Smith, L., Tosomba, O. M. and Grab, D. J. 1995. Trypanosomatid cysteine protease activity may be enhanced by a kininogen-like moiety from host serum. Biochem J 305 (Pt 2), 549-556.

Lorenzo-Morales, J., Ortega-Rivas, A., Foronda, P., Abreu-Acosta, N., Ballart, D., Martinez, E. and Valladares, B. 2005. RNA interference (RNAi) for the silencing of extracellular serine

proteases genes in *Acanthamoeba*: molecular analysis and effect on pathogenecity. Mol Biochem Parasitol 144, 10-15.

Losos, G. J. and Ikede, B. O. 1970. Pathology of experimental trypanosomiasis in the albino rat, rabbit, goat and sheep - A preliminary report. Can J Comp Med 34, 209-212.

Lubega, G. W., Byarugaba, D. K. and Prichard, R. K. 2002. Immunization with a tubulin-rich preparation from *Trypanosoma brucei* confers broad protection against African trypanosomosis. Exp Parasitol 102, 9-22.

Luckins, A. G. 1992. Methods for diagnosis of trypanosomiasis in livestock. In: Branckaert, R. D. S. (Ed.) Trypanosomiasis and tsetse - Africa's disease challenge FAO, Rome.

Lun, Z. R. and Desser, S. S. 1995. Is the broad range of hosts and geographical distribution of *Trypanosoma evansi* attributable to the loss of maxicircle kinetoplast DNA? Parasitol Today 11, 131-133.

Majiwa, P. A., Maina, M., Waitumbi, J. N., Mihok, S. and Zweygarth, E. 1993. *Trypanosoma (Nannomonas) congolense*: molecular characterization of a new genotype from Tsavo, Kenya. Parasitology 106 (Pt 2), 151-162.

Majiwa, P. A. and Webster, P. 1987. A repetitive deoxyribonucleic acid sequence distinguishes *Trypanosoma simiae* from *T. congolense*. Parasitology 95 (Pt 3), 543-558.

Marcello, L. and Barry, J. D. 2007. From silent genes to noisy populations-dialogue between the genotype and phenotypes of antigenic variation. J Eukaryot Microbiol 54, 14-17.

Marcilla, A., De la Rubia, J. E., Sotillo, J., Bernal, D., Carmona, C., Villavicencio, Z., Acosta, D., Tort, J., Bornay, F. J., Esteban, J. G. and Toledo, R. 2008. Leucine aminopeptidase is an immunodominant antigen of *Fasciola hepatica* excretory and secretory products in human infections. Clin Vaccine Immunol 15, 95-100.

Marin-Villa, M., Vargas-Inchaustegui, D. A., Chaves, S. P., Tempone, A. J., Dutra, J. M., Soares, M. J., Ueda-Nakamura, T., Mendonca, S. C., Rossi-Bergmann, B., Soong, L. and Traub-Cseko, Y. M. 2008. The C-terminal extension of *Leishmania pifanoi* amastigote-specific cysteine proteinase Lpcys2: a putative function in macrophage infection. Mol Biochem Parasitol 162, 52-59.

Maser, P., Luscher, A. and Kaminsky, R. 2003. Drug transport and drug resistance in African trypanosomes. Drug Resist Updat 6, 281-290.

Mathew, Z., Knox, T. M. and Miller, C. G. 2000. *Salmonella enterica serovar typhimurium* peptidase B is a leucyl aminopeptidase with specificity for acidic amino acids. J Bacteriol 182, 3383-3393.

Matsui, M., Fowler, J. H. and Walling, L. L. 2006. Leucine aminopeptidases: diversity in structure and function. Biol Chem 387, 1535-1544.

Matthews, K. R., Ellis, J. R. and Paterou, A. 2004. Molecular regulation of the life cycle of African trypanosomes. Trends Parasitol 20, 40-47.

Mattioli, R. C., Faye, J. A. and Büscher, P. 1999. Susceptibility of N'Dama cattle to experimental challenge and cross-species superchallenges with bloodstream forms of *Trypanosoma congolense* and *T. vivax*. Vet Parasitol 86, 83-94.

Mbawa, Z. R., Gumm, I. D., Fish, W. R. and Lonsdale-Eccles, J. D. 1991a. Endopeptidase variations among different life-cycle stages of African trypanosomes. Eur J Biochem 195, 183-190.

Mbawa, Z. R., Gumm, I. D., Shaw, E. and Lonsdale-Eccles, J. D. 1992. Characterisation of a cysteine protease from bloodstream forms of *Trypanosoma congolense*. Eur J Biochem 204, 371-379.

Mbawa, Z. R., Webster, P. and Lonsdale-Eccles, J. D. 1991b. Immunolocalization of a cysteine protease within the lysosomal system of *Trypanosoma congolense*. Eur J Cell Biol 56, 243-250.

McCaman, M. T. and Villarejo, M. R. 1982. Structured and catalytic properties of peptidase N from *Escherichia coli* K-12. Arch Biochem Biophys 213, 384-394.

McCarthy, E., Stack, C., Donnelly, S. M., Doyle, S., Mann, V. H., Brindley, P. J., Stewart, M., Day, T. A., Maule, A. G. and Dalton, J. P. 2004. Leucine aminopeptidase of the human blood flukes, *Schistosoma mansoni* and *Schistosoma japonicum*. Int J Parasitol 34, 703-714.

McGowan, S., Porter, C. J., Lowther, J., Stack, C. M., Golding, S. J., Skinner-Adams, T. S., Trenholme, K. R., Teuscher, F., Donnelly, S. M., Grembecka, J., Mucha, A., Kafarski, P., Degori, R., Buckle, A. M., Gardiner, D. L., Whisstock, J. C. and Dalton, J. P. 2009. Structural basis for the inhibition of the essential *Plasmodium falciparum* M1 neutral aminopeptidase. Proc Natl Acad Sci U S A 106, 2537-2542.

McGrath, M. E., Eakin, A. E., Engel, J. C., McKerrow, J. H., Craik, C. S. and Fletterick, R. J. 1995. The crystal structure of cruzain: a therapeutic target for Chagas' disease. J Mol Biol 247, 251-259.

McKerrow, J. H., Caffrey, C., Kelly, B., Loke, P. and Sajid, M. 2006. Proteases in Parasitic Diseases. Annu. Rev. Pathol. Mech. Dis. 1, 497-536.

McKerrow, J. H., Engel, J. C. and Caffrey, C. R. 1999. Cysteine protease inhibitors as chemotherapy for parasitic infections. Bioorg Med Chem 7, 639-644.

McNamara, J. J., Mohammed, G. and Gibson, W. C. 1994. *Trypanosoma (Nannomonas) godfreyi sp*. nov. from tsetse flies in The Gambia: biological and biochemical characterization. Parasitology 109 (Pt 4), 497-509.

Medina-Acosta, E. and Cross, G. A. 1993. Rapid isolation of DNA from trypanosomatid protozoa using a simple 'mini-prep' procedure. Mol Biochem Parasitol 59, 327-329.

Melville, S. E., Majiwa, P. A. and Tait, A. 2004. The African Trypanosome Genome. In: Maudlin, I., Holmes, P. H., Miles, M. A. (Eds.), The Trypanosomiases. CAB International, Wallingford, pp. 39-57.

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

Menzella, H. G., Gramajo, H. C. and Ceccarelli, E. A. 2002. High recovery of prochymosin from inclusion bodies using controlled air oxidation. Protein Expr Purif 25, 248-255.

Menzies, F. M., Hourez, R., Imarisio, S., Raspe, M., Sadiq, O., Chandraratna, D., O'Kane, C., Rock, K. L., Reits, E., Goldberg, A. L. and Rubinsztein, D. C. 2010. Puromycin-sensitive aminopeptidase protects against aggregation-prone proteins via autophagy. Hum Mol Genet 19, 4573-4586.

Mkhize, P. P. 2003. Epitope mapping of a trypanosomal cysteine proteinase. Biochemistry. University of KwaZulu-Natal, Pietermaritzburg.

Moerschell, R. P., Hosokawa, Y., Tsunasawa, S. and Sherman, F. 1990. The specificities of yeast methionine aminopeptidase and acetylation of amino-terminal methionine in vivo. Processing of altered iso-1-cytochromes c created by oligonucleotide transformation. J Biol Chem 265, 19638-19643.

Monod, M., Lechenne, B., Jousson, O., Grand, D., Zaugg, C., Stocklin, R. and Grouzmann, E. 2005. Aminopeptidases and dipeptidyl-peptidases secreted by the dermatophyte *Trichophyton rubrum*. Microbiology 151, 145-155.

Montagna, G. N., Donelson, J. E. and Frasch, A. C. 2006. Procyclic *Trypanosoma brucei* expresses separate sialidase and trans-sialidase enzymes on its surface membrane. J Biol Chem 281, 33949-33958.

Morales, M. E., Rinaldi, G., Gobert, G. N., Kines, K. J., Tort, J. F. and Brindley, P. J. 2008. RNA interference of *Schistosoma mansoni* cathepsin D, the apical enzyme of the hemoglobin proteolysis cascade. Mol Biochem Parasitol 157, 160-168.

Morrison, L. J., Marcello, L. and McCulloch, R. 2009. Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. Cell Microbiol 11, 1724-1734.

Morty, R. E., Bulau, P., Pelle, R., Wilk, S. and Abe, K. 2006. Pyroglutamyl peptidase type I from *Trypanosoma brucei*: a new virulence factor from African trypanosomes that de-blocks regulatory peptides in the plasma of infected hosts. Biochem J 394, 635-645.

Morty, R. E. and Morehead, J. 2002. Cloning and characterization of a leucyl aminopeptidase from three pathogenic Leishmania species. J Biol Chem 277, 26057-26065.

Morty, R. E., Pelle, R., Vadasz, I., Uzcanga, G. L., Seeger, W. and Bubis, J. 2005. Oligopeptidase B from *Trypanosoma evansi*. A parasite peptidase that inactivates atrial natriuretic factor in the bloodstream of infected hosts. J Biol Chem 280, 10925-10937.

Morty, R. E., Troeberg, L., Pike, R. N., Jones, R., Nickel, P., Lonsdale-Eccles, J. D. and Coetzer, T. H. 1998. A trypanosome oligopeptidase as a target for the trypanocidal agents pentamidine, diminazene and suramin. FEBS Lett 433, 251-256.

Mottram, J. C., Coombs, G. H. and Alexander, J. 2004. Cysteine peptidases as virulence factors of *Leishmania*. Curr Opin Microbiol 7, 375-381.

Mottram, J. C., Frame, M. J., Brooks, D. R., Tetley, L., Hutchison, J. E., Souza, A. E. and Coombs, G. H. 1997. The multiple cpb cysteine proteinase genes of *Leishmania mexicana* encode isoenzymes that differ in their stage regulation and substrate preferences. J Biol Chem 272, 14285-14293.

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

Mottram, J. C., Souza, A. E., Hutchison, J. E., Carter, R., Frame, M. J. and Coombs, G. H. 1996. Evidence from disruption of the *lmcpb* gene array of *Leishmania mexicana* that cysteine proteinases are virulence factors. Proc Natl Acad Sci U S A 93, 6008-6013.

Mucha, A., Drag, M., Dalton, J. P. and Kafarski, P. 2010. Metallo-aminopeptidase inhibitors. Biochimie 92, 1509-1529.

Munday, J. C., McLuskey, K., Brown, E., Coombs, G. H. and Mottram, J. C. 2010. Oligopeptidase B deficient mutants of *Leishmania major*. Mol Biochem Parasitol *In press*.

Mundodi, V., Somanna, A., Farrell, P. J. and Gedamu, L. 2002. Genomic organization and functional expression of differentially regulated cysteine protease genes of *Leishmania donovani* complex. Gene 282, 257-265.

Murray, M., d'Ieteren, G. D. and Teale, A. J. 2004. Trypanotolerance. In: Maudlin, I., Holmes, P. H., Miles, M. A. (Eds.), The Trypanosomiases. CAB International, Wallingford, pp. 461-477.

Murray, M., Murray, P. K. and McIntyre, W. I. 1977. An improved parasitological technique for the diagnosis of African trypanosomiasis. Trans R Soc Trop Med Hyg 71, 325-326.

Na, B. K., Kang, J. M., Cheun, H. I., Cho, S. H., Moon, S. U., Kim, T. S. and Sohn, W. M. 2009. Cryptopain-1, a cysteine protease of *Cryptosporidium parvum*, does not require the pro-domain for folding. Parasitology 136, 149-157.

Naessens, J. 2006. Bovine trypanotolerance: A natural ability to prevent severe anaemia and haemophagocytic syndrome? Int J Parasitol 36, 521-528.

Naessens, J., Teale, A. J. and Sileghem, M. 2002. Identification of mechanisms of natural resistance to African trypanosomiasis in cattle. Vet Immunol Immunopathol 87, 187-194.

Nantulya, V. M. 1990. Trypanosomiasis in domestic animals: the problems of diagnosis. Rev Sci Tech 9, 357-367.

Naula, C., Parsons, M. and Mottram, J. C. 2005. Protein kinases as drug targets in trypanosomes and Leishmania. Biochim Biophys Acta 1754, 151-159.

Ngo, H., Tschudi, C., Gull, K. and Ullu, E. 1998. Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. Proc Natl Acad Sci U S A 95, 14687-14692.

Nok, A. J. and Balogun, E. O. 2003. A bloodstream *Trypanosoma congolense* sialidase could be involved in anemia during experimental trypanosomiasis. J Biochem 133, 725-730.

North, M. J., Mottram, J. C. and Coombs, G. H. 1990. Cysteine proteinases of parasitic protozoa. Parasitol Today 6, 270-275.

O'Farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250, 4007-4021.

Oberholzer, M., Lopez, M. A., McLelland, B. T. and Hill, K. L. 2010. Social motility in African trypanosomes. PLoS Pathog 6, e1000739.

Osório, A. L., Madruga, C. R., Desquesnes, M., Soares, C. O., Ribeiro, L. R. and Costa, S. C. 2008. *Trypanosoma* (Duttonella) *vivax*: its biology, epidemiology, pathogenesis, and introduction in the New World - a review. Mem Inst Oswaldo Cruz 103, 1-13.

Paris, J., Murray, M. and McOdimba, F. 1982. A comparative evaluation of the parasitological techniques currently available for the diagnosis of African trypanosomiasis in cattle. Acta Trop 39, 307-316.

Pays, E. 2005. Regulation of antigen gene expression in *Trypanosoma brucei*. Trends Parasitol 21, 517-520.

Piacenza, L., Acosta, D., Basmadjian, I., Dalton, J. P. and Carmona, C. 1999. Vaccination with cathepsin L proteinases and with leucine aminopeptidase induces high levels of protection against fascioliasis in sheep. Infect Immun 67, 1954-1961.

Pike, R. and Dennison, C. 1989. A high yield method for the isolation of sheep's liver cathepsin L. Prep Biochem 19, 231-245.

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

Pils, B. and Schultz, J. 2004a. Evolution of the multifunctional protein tyrosine phosphatase family. Mol Biol Evol 21, 625-631.

Pils, B. and Schultz, J. 2004b. Inactive enzyme-homologues find new function in regulatory processes. J Mol Biol 340, 399-404.

Playfair, J. H. 1991. Toxins, cytokines, immunity and pathology. Immunol Lett 30, 145-148.

Playfair, J. H., Taverne, J., Bate, C. A. and de Souza, J. B. 1990. The malaria vaccine: anti-parasite or anti-disease? Immunol Today 11, 25-27.

Pokharel, D. R. and Rathaur, S. 2008. Purification and characterization of a leucine aminopeptidase from the bovine filarial parasite *Setaria cervi*. Acta Trop 106, 1-8.

Polson, A., Coetzer, T., Kruger, J., von Maltzahn, E. and van der Merwe, K. J. 1985. Improvements in the isolation of IgY from the yolks of eggs laid by immunized hens. Immunol Invest 14, 323-327.

Price, V. L. and Kieny, M. P. 2001. Vaccines for parasitic diseases. Curr Drug Targets 1, 315-324.

Putrianti, E. D., Schmidt-Christensen, A., Arnold, I., Heussler, V. T., Matuschewski, K. and Silvie, O. 2010. The *Plasmodium* serine-type SERA proteases display distinct expression patterns and non-essential in vivo roles during life cycle progression of the malaria parasite. Cell Microbiol 12, 725-739.

Quintas-Granados, L. I., Orozco, E., Brieba, L. G., Arroyo, R. and Ortega-Lopez, J. 2009. Purification, refolding and autoactivation of the recombinant cysteine proteinase EhCP112 from *Entamoeba histolytica*. Protein Expr Purif 63, 26-32.

Rafati, S., Kariminia, A., Seyde-Eslami, S., Narimani, M., Taheri, T. and Lebbatard, M. 2002. Recombinant cysteine proteinases-based vaccines against *Leishmania major* in BALB/c mice: the partial protection relies on interferon gamma producing CD8(+) T lymphocyte activation. Vaccine 20, 2439-2447.

Rafati, S., Zahedifard, F., Azari, M. K., Taslimi, Y. and Taheri, T. 2008. *Leishmania infantum*: prime boost vaccination with C-terminal extension of cysteine proteinase type I displays both type 1 and 2 immune signatures in BALB/c mice. Exp Parasitol 118, 393-401.

Ralston, K. S. and Hill, K. L. 2008. The flagellum of *Trypanosoma brucei:* new tricks from an old dog. Int J Parasitol 38, 869-884.

Ramos, A. M., Duschak, V. G., Gerez de Burgos, N. M., Barboza, M., Remedi, M. S., Vides, M. A. and Chiabrando, G. A. 2002. *Trypanosoma cruzi*: cruzipain and membrane-bound cysteine

proteinase isoform(s) interacts with human alpha(2)-macroglobulin and pregnancy zone protein. Exp Parasitol 100, 121-130.

Rasooly, R. and Balaban, N. 2004. Trypanosome microtubule-associated protein p15 as a vaccine for the prevention of African sleeping sickness. Vaccine 22, 1007-1015.

Rassi, A., Jr., Rassi, A. and Marin-Neto, J. A. 2010. Chagas disease. Lancet 375, 1388-1402.

Rawlings, N. D. and Barrett, A. J. 1993. Evolutionary families of peptidases. Biochem J 290 205-218.

Rawlings, N. D., Barrett, A. J. and Bateman, A. 2010. MEROPS: the peptidase database. Nucleic Acids Res 38, D227-233.

Rayaisse, J. B., Tirados, I., Kaba, D., Dewhirst, S. Y., Logan, J. G., Diarrassouba, A., Salou, E., Omolo, M. O., Solano, P., Lehane, M. J., Pickett, J. A., Vale, G. A., Torr, S. J. and Esterhuizen, J. 2010. Prospects for the development of odour baits to control the tsetse flies *Glossina tachinoides* and *G. palpalis s.l*. PLoS Negl Trop Dis 4, e632.

Reis, F. C., Costa, T. F., Sulea, T., Mezzetti, A., Scharfstein, J., Brömme, D., Ménard, R. and Lima, A. P. 2007. The propeptide of cruzipain-a potent selective inhibitor of the trypanosomal enzymes cruzipain and brucipain, and of the human enzyme cathepsin F. Febs J 274, 1224-1234.

Reszka, N., Rijsewijk, F. A., Zelnik, V., Moskwa, B. and Bienkowska-Szewczyk, K. 2007. *Haemonchus contortus*: characterization of the baculovirus expressed form of aminopeptidase H11. Exp Parasitol 117, 208-213.

Rinaldi, G., Morales, M. E., Alrefaei, Y. N., Cancela, M., Castillo, E., Dalton, J. P., Tort, J. F. and Brindley, P. J. 2009. RNA interference targeting leucine aminopeptidase blocks hatching of *Schistosoma mansoni* eggs. Mol Biochem Parasitol 167, 118-126.

Robinson, M. W., Dalton, J. P. and Donnelly, S. 2008. Helminth pathogen cathepsin proteases: it's a family affair. Trends Biochem Sci 33, 601-608.

Roche, L., Tort, J. and Dalton, J. P. 1999. The propeptide of *Fasciola hepatica* cathepsin L is a potent and selective inhibitor of the mature enzyme. Mol Biochem Parasitol 98, 271-277.

Rogers, D. J. and Randolph, S. E. 2002. A response to the aim of eradicating tsetse from Africa. Trends Parasitol 18, 534-536.

Rosenthal, P. J. 2004. Cysteine proteases of malaria parasites. Int J Parasitol 34, 1489-1499.

Roy, N., Nageshan, R. K., Pallavi, R., Chakravarthy, H., Chandran, S., Kumar, R., Gupta, A. K., Singh, R. K., Yadav, S. C. and Tatu, U. 2010. Proteomics of *Trypanosoma evansi* infection in rodents. PLoS One 5, e9796.

Rozema, D. and Gellman, S. H. 1996. Artificial chaperone-assisted refolding of carbonic anhydrase B. J Biol Chem 271, 3478-3487.

Sajid, M. and McKerrow, J. H. 2002. Cysteine proteases of parasitic organisms. Mol Biochem Parasitol 120, 1-21.

Salvesen, G. S. and Nagase, H. 2001. Inhibition of proteolytic enzymes. In: Benyon, R., Bond, J. S. (Eds.), Proteolytic Enzymes. Oxford University Press, Oxford, pp. 105-130.

Samarasinghe, B., Knox, D. P. and Britton, C. 2010. Factors affecting susceptibility to RNA interference in *Haemonchus contortus* and in vivo silencing of an H11 aminopeptidase gene. Int J Parasitol *In press*.

Sanger, F. and Coulson, A. R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 94, 441-448.

Santos, A. N., Langner, J., Herrmann, M. and Riemann, D. 2000. Aminopeptidase N/CD13 is directly linked to signal transduction pathways in monocytes. Cell Immunol 201, 22-32.

Sarkany, Z. and Polgar, L. 2003. The unusual catalytic triad of poliovirus protease 3C. Biochemistry 42, 516-522.

Schechter, I. and Berger, A. 1967. On the size of the active site in proteases. I. Papain. Biochem Biophys Res Commun 27, 157-162.

Schenkman, S., Eichinger, D., Pereira, M. E. and Nussenzweig, V. 1994. Structural and functional properties of *Trypanosoma* trans-sialidase. Annu Rev Microbiol 48, 499-523.

Schofield, L. 2007. Rational approaches to developing an anti-disease vaccine against malaria. Microbes Infect 9, 784-791.

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

Serveau, C., Lalmanach, G., Juliano, M. A., Scharfstein, J., Juliano, L. and Gauthier, F. 1996. Investigation of the substrate specificity of cruzipain, the major cysteine proteinase of *Trypanosoma cruzi*, through the use of cystatin-derived substrates and inhibitors. Biochem J 313 (Pt 3), 951-956.

Sethadavit, M., Meemon, K., Jardim, A., Spithill, T. W. and Sobhon, P. 2009. Identification, expression and immunolocalization of cathepsin B3, a stage-specific antigen expressed by juvenile *Fasciola gigantica*. Acta Trop 112, 164-173.

Sharma, A. 2007. Malarial protease inhibitors: potential new chemotherapeutic agents. Curr Opin Investig Drugs 8, 642-652.

Sharma, R., Gluenz, E., Peacock, L., Gibson, W., Gull, K. and Carrington, M. 2009. The heart of darkness: growth and form of *Trypanosoma brucei* in the tsetse fly. Trends Parasitol 25, 517-524.

Shi, H., Djikeng, A., Mark, T., Wirtz, E., Tschudi, C. and Ullu, E. 2000. Genetic interference in *Trypanosoma brucei* by heritable and inducible double-stranded RNA. RNA 6, 1069-1076.

Sijwali, P. S., Brinen, L. S. and Rosenthal, P. J. 2001. Systematic optimization of expression and refolding of the *Plasmodium falciparum* cysteine protease falcipain-2. Protein Expr Purif 22, 128-134.

Sijwali, P. S., Koo, J., Singh, N. and Rosenthal, P. J. 2006. Gene disruptions demonstrate independent roles for the four falcipain cysteine proteases of *Plasmodium falciparum*. Mol Biochem Parasitol 150, 96-106.

Sijwali, P. S., Shenai, B. R. and Rosenthal, P. J. 2002. Folding of the *Plasmodium falciparum* cysteine protease falcipain-2 is mediated by a chaperone-like peptide and not the prodomain. J Biol Chem 277, 14910-14915.

Singh, S. M. and Panda, A. K. 2005. Solubilization and refolding of bacterial inclusion body proteins. J Biosci Bioeng 99, 303-310.

Skinner-Adams, T. S., Stack, C. M., Trenholme, K. R., Brown, C. L., Grembecka, J., Lowther, J., Mucha, A., Drag, M., Kafarski, P., McGowan, S., Whisstock, J. C., Gardiner, D. L. and Dalton, J. P. 2009. *Plasmodium falciparum* neutral aminopeptidases: new targets for anti-malarials. Trends Biochem Sci 35, 53-61.

Smith, T. S., Graham, M., Munn, E. A., Newton, S. E., Knox, D. P., Coadwell, W. J., McMichael-Phillips, D., Smith, H., Smith, W. D. and Oliver, J. J. 1997. Cloning and characterization of a microsomal aminopeptidase from the intestine of the nematode *Haemonchus contortus*. Biochim Biophys Acta 1338, 295-306.

Smooker, P. M., Jayaraj, R., Pike, R. N. and Spithill, T. W. 2010. Cathepsin B proteases of flukes: the key to facilitating parasite control? Trends Parasitol *In press*.

Song, S. M., Park, J. H., Kim, J., Kim, S. I., Hong, Y. C., Kong, H. H. and Chung, D. I. 2008. Identification and characterization of *Paragonimus westermani* leucine aminopeptidase. Parasitol Int 57, 334-341.

Souza, A. E., Waugh, S., Coombs, G. H. and Mottram, J. C. 1992. Characterization of a multi-copy gene for a major stage-specific cysteine proteinase of *Leishmania mexicana*. FEBS Lett 311, 124-127.

Sow, A., Sidibe, I., Desquesnes, M., Bengaly, Z. and Pangui, L. J. 2006. The application of PCR-ELISA to the detection of *Trypanosoma congolense* type savannah (TCS) in bovine blood samples. Trop Biomed 23, 123-129.

Spinks, D., Shanks, E. J., Cleghorn, L. A., McElroy, S., Jones, D., James, D., Fairlamb, A. H., Frearson, J. A., Wyatt, P. G. and Gilbert, I. H. 2009. Investigation of trypanothione reductase as a drug target in *Trypanosoma brucei*. Chem Med Chem 4, 2060-2069.

Sreedharan, S. K., Verma, C., Caves, L. S., Brocklehurst, S. M., Gharbia, S. E., Shah, H. N. and Brocklehurst, K. 1996. Demonstration that 1-trans-epoxysuccinyl-L-leucylamido-(4-guanidino) butane (E-64) is one of the most effective low Mr inhibitors of trypsin-catalysed hydrolysis. Characterization by kinetic analysis and by energy minimization and molecular dynamics simulation of the E-64-beta-trypsin complex. Biochem J 316 (Pt 3), 777-786.

Stack, C. M., Lowther, J., Cunningham, E., Donnelly, S., Gardiner, D. L., Trenholme, K. R., Skinner-Adams, T. S., Teuscher, F., Grembecka, J., Mucha, A., Kafarski, P., Lua, L., Bell, A. and Dalton, J. P. 2007. Characterization of the *Plasmodium falciparum* M17 leucyl aminopeptidase. A protease involved in amino acid regulation with potential for antimalarial drug development. J Biol Chem 282, 2069-2080.

Stanton, J. D., Rashid, M. B. and Mensa-Wilmot, K. 2002. Cysteine-less glycosylphosphatidylinositol-specific phospholipase C is inhibited competitively by a thiol reagent: evidence for glyco-mimicry by p-chloromercuriphenylsulphonate. Biochem J 366, 281-288.

Stevens, J. R. and Brisse, S. 2004. Systematics of trypanosomes of Medical and Veterinary importance. In: Maudlin, I., Holmes, P. H., Miles, M. A. (Eds.), The Trypanosomiases. CABI Publishing, Oxfordshire, pp. 1-19.

Steverding, D. 2008. The history of African trypanosomiasis. Parasit Vectors 1, 3.

Steverding, D. 2010. The development of drugs for treatment of sleeping sickness: a historical review. Parasit Vectors 3, 15.

Swenerton, R. K., Zhang, S., Sajid, M., Medzihradszky, K. F., Craik, C. S., Kelly, B. L. and McKerrow, J. H. 2010. The Oligopeptidase B of *Leishmania regulates* parasite enolase and immune evasion. J Biol Chem *In press*.

Swinkels, B. W., Evers, R. and Borst, P. 1988. The topogenic signal of the glycosomal (microbody) phosphoglycerate kinase of *Crithidia fasciculata* resides in a carboxy-terminal extension. Embo J 7, 1159-1165.

Szoor, B. 2010. Trypanosomatid protein phosphatases. Mol Biochem Parasitol 173, 53-63.

Tanowitz, H. B., Kirchhoff, L. V., Simon, D., Morris, S. A., Weiss, L. M. and Wittner, M. 1992. Chagas' disease. Clin Microbiol Rev 5, 400-419.

Taylor, A. 1993. Aminopeptidases: structure and function. FASEB J 7, 290-298.

Taylor, J. E. and Rudenko, G. 2006. Switching trypanosome coats: what's in the wardrobe? Trends Genet 22, 614-620.

Taylor, K. A. 1998. Immune responses of cattle to African trypanosomes: protective or pathogenic? Int J Parasitol 28, 219-240.

Taylor, K. A. and Authié, E. 2004. Pathogenesis of African Trypanosomiasis. In: Maudlin, I., Holmes, P. H., Miles, M. A. (Eds.), The Trypanosomiases. CAB International, Wallingford, pp. 331-353.

Tizard, I., Nielsen, K. H., Seed, J. R. and Hall, J. E. 1978. Biologically active products from African Trypanosomes. Microbiol Rev 42, 664-681.

Tomás, A. M. and Kelly, J. M. 1996. Stage-regulated expression of cruzipain, the major cysteine protease of *Trypanosoma cruzi* is independent of the level of RNA1. Mol Biochem Parasitol 76, 91-103.

Torr, S. J., Hargrove, J. W. and Vale, G. A. 2005. Towards a rational policy for dealing with tsetse. Trends Parasitol 21, 537-541.

Torrie, L. S., Wyllie, S., Spinks, D., Oza, S. L., Thompson, S., Harrison, J. R., Gilbert, I. H., Wyatt, P. G., Fairlamb, A. H. and Frearson, J. A. 2009. Chemical validation of trypanothione

synthetase: a potential drug target for human trypanosomiasis. J Biol Chem 284, 36137-36145.

Towbin, H., Staehelin, T. and Gordon, J. 1992. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979. Biotechnology 24, 145-149.

Troeberg, L., Chen, X., Flaherty, T. M., Morty, R. E., Cheng, M., Hua, H., Springer, C., McKerrow, J. H., Kenyon, G. L., Lonsdale-Eccles, J. D., Coetzer, T. H. and Cohen, F. E. 2000. Chalcone, acyl hydrazide, and related amides kill cultured *Trypanosoma brucei brucei*. Mol Med 6, 660-669.

Troeberg, L., Morty, R. E., Pike, R. N., Lonsdale-Eccles, J. D., Palmer, J. T., McKerrow, J. H. and Coetzer, T. H. 1999. Cysteine proteinase inhibitors kill cultured bloodstream forms of *Trypanosoma brucei brucei*. Exp Parasitol 91, 349-355.

Uilenberg, G. 1998. A field guide for the diagnosis, treatment and prevention of African animal trypanosomosis FAO Corporate Document Repository, Rome.

Ullu, E., Tschudi, C. and Chakraborty, T. 2004. RNA interference in protozoan parasites. Cell Microbiol 6, 509-519.

Van den Bossche, P. 2001. Some general aspects of the distribution and epidemiology of bovine trypanosomosis in southern Africa. Int J Parasitol 31, 592-598.

Van der Ploeg, L. H., Cornelissen, A. W., Barry, J. D. and Borst, P. 1984. Chromosomes of kinetoplastida. Embo J 3, 3109-3115.

van Deursen, F. J., Thornton, D. J. and Matthews, K. R. 2003. A reproducible protocol for analysis of the proteome of *Trypanosoma brucei* by 2-dimensional gel electrophoresis. Mol Biochem Parasitol 128, 107-110.

van Regenmortel, M. H. V. and Muller, S. 1999. Molecular dissection of protein antigens and the prediction of epitopes. In: van der Vliet, P. C. (Ed.) Synthetic peptides as antigens. Elsevier, Amsterdam, pp. 62-63.

Vazeux, G., Iturrioz, X., Corvol, P. and Llorens-Cortes, C. 1998. A glutamate residue contributes to the exopeptidase specificity in aminopeptidase A. Biochem J 334 407-413.

Vercruysse, J., Knox, D. P., Schetters, T. P. and Willadsen, P. 2004. Veterinary parasitic vaccines: pitfalls and future directions. Trends Parasitol 20, 488-492.

Verlinde, C. L., Hannaert, V., Blonski, C., Willson, M., Perie, J. J., Fothergill-Gilmore, L. A., Opperdoes, F. R., Gelb, M. H., Hol, W. G. and Michels, P. A. 2001. Glycolysis as a target for the design of new anti-trypanosome drugs. Drug Resist Updat 4, 50-65.

Vickerman, K. 1969a. The fine structure of *Trypanosoma congolense* in its bloodstream phase. J Protozool 16, 54-69.

Vickerman, K. 1969b. On the surface coat and flagellar adhesion in trypanosomes. J Cell Sci 5, 163-193.

Vickerman, K. 1986. Parasitology. Clandestine sex in trypanosomes. Nature 322, 113-114.

Vickerman, K., Tetley, L., Hendry, K. A. and Turner, C. M. 1988. Biology of African trypanosomes in the tsetse fly. Biol Cell 64, 109-119.

Vincendeau, P., Lesthelle, S., Bertazzo, A., Okomo-Assoumou, M. C., Allegri, G. and Costa, C. V. L. 1999. Important of tryptophan metabolism in trypanosomiasis. In: Huether, G., Kochen, W., Simat, T. J., Steinhart, H. (Eds.), Tryptophan, serotonin and melatonin: Basic aspects and applications. Springer, Hamburg, pp. 525-531.

Vreysen, M. J. 2001. Principles of area-wide integrated tsetse fly control using the sterile insect technique. Med Trop (Mars) 61, 397-411.

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. and Feldmann, H. U. 2000. *Glossina austeni* (Diptera: Glossinidae) eradicated on the island of Unguja, Zanzibar, using the sterile insect technique. J Econ Entomol 93, 123-135.

Wainszelbaum, M., Isola, E., Wilkowsky, S., Cannata, J. J., Florin-Christensen, J. and Florin-Christensen, M. 2001. Lysosomal phospholipase A1 in *Trypanosoma cruzi*: an enzyme with a possible role in the pathogenesis of Chagas' disease. Biochem J 355, 765-770.

Wang, Z., Morris, J. C., Drew, M. E. and Englund, P. T. 2000. Inhibition of *Trypanosoma brucei* gene expression by RNA interference using an integratable vector with opposing T7 promoters. J Biol Chem 275, 40174-40179.

Wellde, B., Lotzsch, R., Deindl, G., Sadun, E., Williams, J. and Warui, G. 1974. *Trypanosoma congolense*. I. Clinical observations of experimentally infected cattle. Exp Parasitol 36, 6-19.

Welling, G. W., Weijer, W. J., van der Zee, R. and Welling-Wester, S. 1985. Prediction of sequential antigenic regions in proteins. FEBS Lett 188, 215-218.

Wilson, A. J. 1969. Value of the indirect fluorescent antibody test as a serological aid to diagnosis of *Glossina*-transmitted bovine trypanosomiasis. Trop. Anita. Hlth Prod 1, 89-95.

Wilson, L. R., Good, R. T., Panaccio, M., Wijffels, G. L., Sandeman, R. M. and Spithill, T. W. 1998. *Fasciola hepatica*: characterization and cloning of the major cathepsin B protease secreted by newly excysted juvenile liver fluke. Exp Parasitol 88, 85-94.

Wint, W. and Rogers, D. 1999. Predicted tsetse distributions. Environmental Research Group Oxford (ERGO).

Wirtz, E., Leal, S., Ochatt, C. and Cross, G. A. 1999. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. Mol Biochem Parasitol 99, 89-101.

Wrightsman, R. A. and Manning, J. E. 2000. Paraflagellar rod proteins administered with alum and IL-12 or recombinant adenovirus expressing IL-12 generates antigen-specific responses and protective immunity in mice against *Trypanosoma cruzi*. Vaccine 18, 1419-1427.

Wu, S. and Letchworth, G. J. 2004. High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol. Biotechniques 36, 152-154.

Zadeh-Vakili, A., Taheri, T., Taslimi, Y., Doustdari, F., Salmanian, A. H. and Rafati, S. 2004. Immunization with the hybrid protein vaccine, consisting of *Leishmania major* cysteine proteinases Type I (CPB) and Type II (CPA), partially protects against leishmaniasis. Vaccine 22, 1930-1940.

Zambrano-Villa, S., Rosales-Borjas, D., Carrero, J. C. and Ortiz-Ortiz, L. 2002. How protozoan parasites evade the immune response. Trends Parasitol 18, 272-278.