GENE DISRUPTION OF *Tco*CATL (CONGOPAIN) AND OLIGOPEPTIDASE B, PATHOGENIC FACTORS OF AFRICAN TRYPANOSOMES.

By

RICHARD THIGA KANGETHE

(M.Sc., University of Nairobi)

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 2011

DECLARATION – PLAGIARISM

I, Richard T. Kangethe, 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 from March 2007 to December 2011 and the University of Bordeaux 2, France from April-May 2007, October-November 2008 and October-November 2009, under the supervision of Prof. Theresa H.T. Coetzer.

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: Richard T. Kangethe

_____

Supervisor: Prof. Theresa H.T. Coetzer

ABSTRACT

African trypanosomosis is a parasitic disease in man and animals caused by protozoan parasites of the genus *Trypanosoma*. *T. congolense, T. vivax* and *T. brucei brucei* cause nagana in cattle. The variable nature of the parasite surface coat has hindered the development of an effective vaccine. An option for developing vaccines and chemotherapeutic agents against trypanosomosis is to target pathogenic factors released by the parasite during infection, namely an "anti-disease" approach. Two pathogenic factors released during infection are oligopeptidase B (OPB) and *Tco*CATL (congopain). *Tco*CATL, a major lysosomal cysteine peptidase, is a member of the papain family C1 cysteine peptidases. RNA interference (RNAi) was used to down-regulate the expression of *Tco*CATL in *T. congolense* IL3000 TRUM183:29-13 parasites *in vivo* during mouse infections. *Tco*CATL RNAi was monitored in infected mouse blood by comparing the hydrolysis of Z-Phe-Arg-AMC and parasitaemia between mice in which RNAi was induced and control mice. Mice infected with parasites induced for *Tco*CATL RNAi had lower parasitaemia when compared to control mice. An attempt was also made at deleting the entire CATL gene array in both *T. congolense* IL3000 and *T. brucei* 427 Lister strains. The second pathogenic factor studied, OPB, is a cytosolic trypanosomal peptidase that hydrolyses peptides smaller than 30 amino acid residues, C-terminal to basic residues. In order to evaluate the role that OPB play during disease, RNAi was also applied to knock-down the expression levels of OPB in *T. brucei* T7T and *T. congolense* IL3000 TRUM183:29-13 strains (*Tb*OPB and *Tco*OPB respectively). Oligopeptidase B null mutant strains (Δ*opb*) were also generated in *T. brucei brucei* Lister 427. An attempt was also made to generate OPB null mutants in *T. congolense* IL3000 parasites. Western blot analysis of the knock-down experiments using chicken anti-*Tco*OPB peptide IgY showed that only *Tb*OPB levels were reduced in *T. brucei* T7T parasites induced for RNAi when compared to *Tco*OPB RNAi induced cultures. Quantitative assessment of a fourteen day induction experiment for OPB RNAi in *T. brucei* showed an 87% reduction in *Tb*OPB levels when compared to levels on day one. There was no growth effect observed in *T. brucei* parasites cultured *in vitro* and induced for *Tb*OPB RNAi. It was concluded that *Tb*OPB is not necessary for the *in vitro* survival of *T. brucei* parasites, thus making the generation of OPB null mutants possible. Δ*opb T. brucei* parasites were successfully generated and grew normally *in vitro* and were as virulent as wild type strains during infection in mice. Immunohistopatholgy of infected mouse testes revealed Δ*opb* parasites in extra vascular regions showing that *T. brucei* OPB (*Tb*OPB) is not involved in assisting *T. brucei* parasites to cross microvascular endothelial cells. Gelatin gel analysis of Δ*opb* null mutants

and wild type strains showed an increase in cysteine peptidase activity. Enzymatic activity assays were carried out to identify how closely related oligopeptidases are affected by knocking out *TbOPB*, and a significant increase of *T. brucei* prolyl oligopeptidase (*Tb*POP) activity was observed. However, western blot analysis did not show any increase of *Tb*POP protein levels in Δ*opb* parasites, suggesting that either *Tb*OPB is responsible for generating an endogenous inhibitor for *Tb*POP or that another POP-like enzyme might compensate for a loss in OPB activity in Δ*opb* null mutants. This study made a significant contribution to an understanding of the interplay between different trypanosomal peptidases that are important pathogenic factors in trypanosomosis. It highlights the need to simultaneously target several trypanosomal peptidases to develop an effective vaccine or chemotherapeutic agents for African animal trypanosomosis.

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to the following people and institutions: My supervisor, Prof. Theresa Coetzer, for her guidance, patience and assistance throughout this study and for her critical reading of this thesis.

Dr Alain Boulangé for introducing me to trypanosomosis research and offering his assistance over the course of my studies.

Prof. Théo Baltz and Virginie Coustou for their assistance and expertise at 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.

My colleagues Phillia Vukea, Davita Pillay, Hlumani Ndlovu, Lorelle Bizaaré, Cara Bartlett, Perina Vather, Mayuri Jugmohan, Sabelo Hadebe, Ike Achilonu, Jackie Viljoen, Laurelle Jackson, Kayleen O'brien, Kyle Goetsch, Celia Snyman and Bridgette Cumming for their technical assistance when required and constructive advice.

And lastly to my parents, Mr. and Mrs. Kangethe and my siblings, it has been a long journey.

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).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

2x YT	yeast tryptone media
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride
AMC	aminomethyl coumarin
BLASTN	basic local alignment search tool - nucleotide
BSA	bovine serum albumin
CP(s)	cysteine peptidase(s)
DEAE	diethyl-aminoethyl
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
HRPO	horse radish peroxidase
IFA	immunofluorescence assay
IFNγ	interferon gamma
Ig	immunoglobulin
ILRI	International Livestock Research Institute
mAb	monoclonal antibody
OD600	optical density at 600 nm
OPB	oligopeptidase B
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
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
sVSG	soluble variant surface glycoprotein
TBS	tris buffered saline

CATB	cathepsin-B
CATL	cathepsin-L
TNFα	tumour necrosis factor alpha
VAT	variant antigen type
VSG	variant surface glycoprotein
WT	wild type
Z	benzyloxycarbonyl

CHAPTER 1

INTRODUCTION

1.1 GENERAL OVERVIEW

African trypanosomosis is a parasitic disease in man and animals caused by protozoan parasites of the genus *Trypanosoma*. On the African continent, these parasites are mainly transmitted by the tsetse fly (*Glossina spp.*). This disease presents a major obstacle to development in sub-Saharan Africa and is responsible for the loss of human life, depletion of livestock and underutilisation of arable land (Swallow, 2000; McDermott and Coleman, 2001; Antoine-Moussiaux *et al.*, 2009). The disease occurs as far south as the northern part of KwaZulu-Natal (Kappmeier *et al.*, 1998), and recent studies identifying infections in cattle makes a study of the disease particularly relevant in the local context (Mamabolo *et al.*, 2009). In cattle, the disease is mainly caused by *T. congolense* and *T. vivax*, while *T. b. brucei* is less pathogenic. Various approaches have been used to manage the spread of the disease in livestock, mainly through vector control (Brown, 2008) and chemotherapy (Delespaux and de Koning, 2007), with various degrees of success and at a very high financial cost. Vaccine development presents an attractive alternative for the control of this disease in cattle. Initial vaccine targets included surface antigens of the parasite (Taylor and Mertens, 1999). However, the life cycle of African trypanosomes is exclusively extracellular and parasites survive mammalian host immune responses by constantly changing their surface antigens through a mechanism known as antigenic variation (Taylor and Rudenko, 2006), thereby hindering the development of vaccine formulations (Magez and Radwanska, 2009). More recently, an “anti-disease” vaccine approach has been pursued targeting the invariant disease-causing factors released by the parasite rather than the infecting parasite itself (Schofield *et al.*, 1993; Authié *et al.*, 2001; Schofield, 2007). Particularly relevant are trypanosome proteolytic enzymes such as proteinases and peptidases e.g. congopain or *Tco*CATL (Authié *et al.*, 1992) and oligopeptidase B or OPB (Coetzer *et al.*, 2008). These peptidases provide a more attractive target than surface antigens as they are invariant. In addition, they can also be chemotherapeutic targets and pave the way for the development of new drugs.

The inclusion of pathogenic factors in an effective sub-unit vaccine or as drug targets requires studies that elucidate their function in parasite physiology and the role that they may play during infection. Reverse genetics is the tool of choice for this type of study. Reverse genetics

is a technology that enables researchers to study the function of a gene by either down-regulating its messenger RNA (Ullu *et al.*, 2002), or by deleting it completely from the genome of the organism under study (Xu *et al.*, 2009). The elimination of mRNA can be achieved by inducible RNA interference (RNAi), whereas the deletion of a gene is effected by gene knock-down experiments. RNAi has been used as a tool in a variety of organisms including trypanosomes (Ullu *et al.*, 2004). The RNAi phenomenon exists naturally in several eukaryotes and plays a role in defence against viral infections and in the regulation of gene expression. It is defined as the mechanism through which gene-specific, double stranded RNA (dsRNA) triggers the degradation of homologous transcripts (Fire *et al.*, 1998). This process is mediated by a natural biological system that transforms the dsRNA into shorter versions called small interfering RNAs (siRNA) (Shi *et al.*, 2006a). The siRNAs then drive gene silencing by sequence specific messenger RNA (mRNA) degradation thereby leading to a reduction in protein synthesis from the targeted gene. Gene knock-out experiments involve the complete deletion of a gene from an organism by homologous recombination events (Xu *et al.*, 2009). This involves constructing DNA vectors that carry sequences homologous to the regions flanking the gene to be studied that will excise the gene from the genome of the organism by replacing it with a drug resistance gene. This then allows for selection of clones that do not carry the gene of interest any longer by their ability to grow in media containing the drug (Ommen *et al.*, 2009).

This study used RNAi and gene deletion to investigate the function of two trypanosome peptidases with documented pathogenic effects during infection, namely *Tco*CATL and OPB. Inducible gene silencing using RNAi and the deletion of genes allows access to two types of information. Firstly, whether the targeted genes are necessary for the survival of the parasites, and if the gene product can therefore be targeted for chemotherapy and secondly, whether the targeted genes are indeed important pathogenic factors, by *in vivo* infection studies comparing the effect of gene knock-out clones versus wild type parasites.

1.2 TRYPANOSOMES AND TRYPANOSOMOSIS

African trypanosomosis is a mammalian parasitic disease caused by protozoans of the genus *Trypanosoma,* which are transmitted by the tsetse fly (*Glossina spp.*). On the African continent, tsetse flies are distributed over a land mass area of approximately 10 million km^2 and present a major obstacle to human and livestock development, and the utilisation of

arable land (Kuzoe, 1993). There are several species of trypanosomes that affect both humans and animals. In humans, the most important is *T. brucei* which has two sub species that are the main cause of sleeping sickness i.e. *T. brucei rhodesiense* and *T. brucei gambiense* (El-Sayed *et al.*, 2000). African trypanosomes implicated in cattle disease are *T. congolense* and *T. vivax* and to a lesser extent *T. b. brucei*. Trypanosomosis in cattle is also referred to as *nagana* (Taylor, 1998). There are about 172 million head of cattle kept in the 37 countries in sub-Saharan Africa that have tsetse infestation, covering an area of approximately 8.7 million km^2 (Swallow, 2000). Annual losses due to the disease are estimated to be around 1.3 to 5 billion US$ (Kristjanson *et al.*, 1999; McDermott and Coleman, 2001). In South Africa, trypanosomosis in cattle was identified in the late 1800s (Taylor, 1998; Brown, 2008). In 1894 while working in Zululand, Sir David Bruce demonstrated the transmission of a protozoan parasite which was later named *T. b. brucei,* by tsetse flies in cattle. He was also able to show that wild herbivores were asymptomatic following trypanosome infections and were likely reservoirs for the disease in livestock (Brown, 2008). Reports of trypanosome infections in South African cattle have been as recent as 2009 in Northern parts of KwaZulu-Natal close to the Hluhluwe-Umfolozi game reserve (Mamabolo *et al.*, 2009). The cattle were found to be infected with *T. congolense* and *T. vivax*.

1.2.1 Classification of Trypanosomes

The genus *Trypanosoma* falls under the class Zoomastigophora where members of the class are classified by the number of flagellar they possess (Vickerman, 1982; Vickerman, 2000). Members of the order Kinetoplastida possess one or two flagellar with trypanosomes well represented due to their economic importance. *Leishmania* and *Trypanosoma* are the two economically most important genera in this order. Members of the genus *Trypanosoma* are divided into two sections, stercoraria and salivaria, depending on the position of their infective forms in the vector host gut during development. Members of the section stercoraria e.g. *T. cruzi*, develop into metacyclics at the posterior end of the digestive tract of the vector. By contrast, members of section salivaria e.g. *T. brucei*, develop at the anterior end and are transmitted to the mammalian host via saliva. Fig.1.1 gives a comprehensive classification of the Trypanosoma genus.

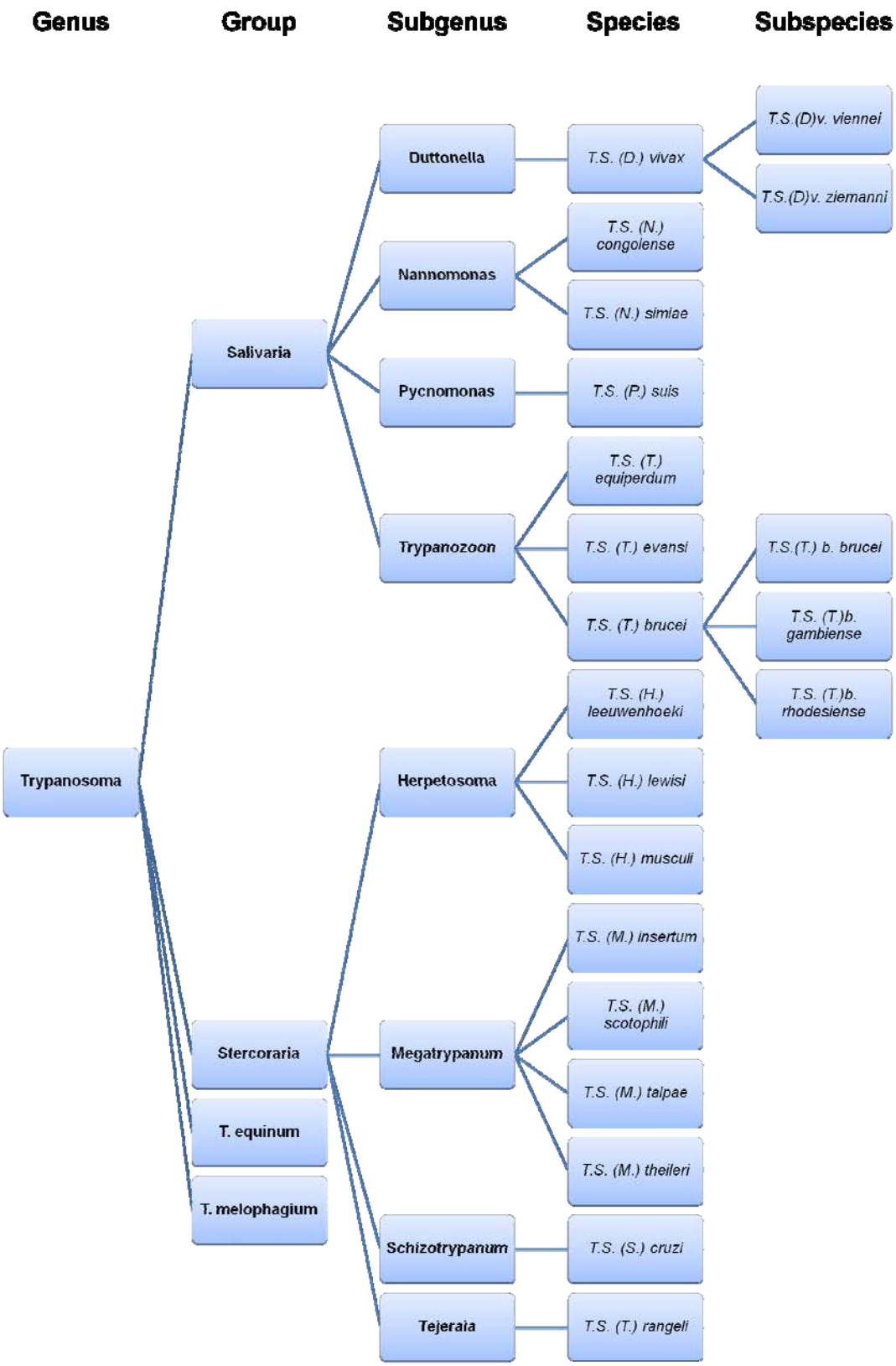


Figure 1.1 Classification of trypanosomes. [Adapted from Vickerman (2000)]

1.2.2 Biology and life cycle of trypanosomes

Trypanosomes infect many different hosts (Table 1.1) and use a variety of vectors to transmit the disease between the primary hosts. For example, *T. brucei* is transmitted between mammalian hosts through a tsetse fly vector. In contrast, *T. equiperdum* is sexually transmitted between horses and spends no life cycle stage in insects. *T. rangeli* and *T. cruzi* are transmitted between hosts through triatomine vectors such as the bed bug. Trypanosomes are mononuclear and possess a large mitochondrion that contains a kinetoplast (Fig.1.2). The kinetoplast contains both maxicircles and minicircles of DNA, where mRNA transcribed in the maxi-circles is modified by guide-RNA from the minicircles (Ryan and Englund, 1989). This is thought to regulate gene expression and play a role in mitochondrial metabolism at different stages in the life cycle of a trypanosome (Simpson, 1987). Trypanosomes have a single flagellum that arises at the posterior end from the basal body in the floor of the flagellar pocket, with an extending paraxial rod (Noble and Noble, 1982). Flagellar functions involve motility, attachment to host surfaces, and a role in morphogenesis and cell division (Kohl and Bastin, 2005). The flagellum emerges from the flagellar pocket, which is an important component involved in immune evasion by the internalisation of VSG-antibody complexes through the action of phospholipase-C (Field *et al.*, 2009). The flagellar pocket is also the principal site at which host macromolecules are internalised.

The entire life cycle of African trypanosomes is extracellular (Fig.1.3). In *T. brucei*, metacyclic trypomastigotes are introduced to the mammalian host by the tsetse fly while taking a blood meal where they rapidly transform into antigenically variable bloodstream forms and divide by binary fission causing a build-up of metabolic and cell waste that forms a chancre at the site of the bite (Baral, 2010). From the chancre, bloodstream trypomastigotes enter the lymphatic system and pass into the bloodstream where they are carried to other sites throughout the body, reaching other blood fluids (e.g. lymph and spinal fluid). Bloodstream parasites continue to replicate transforming from long slender forms to intermediate and eventually into non-dividing short stumpy forms that are once again taken up by the tsetse fly during a bloodmeal on an infected host (Vickerman, 1985). In the fly's midgut, the parasites transform into procyclic trypomastigotes. Fourteen to 21 days later, the procyclic forms migrate into the proventricular region of the gut and multiply asexually into epimastigotes. The epimastigotes then reach the fly's salivary glands where they continue dividing into metacyclics, ready to infect another host. The whole cycle in the fly takes approximately three weeks to progress

(Noble and Noble, 1982). In *T. congolense*, metacyclic forms develop in the mouth parts as opposed to the salivary glands in *T. brucei*, whereas in *T. vivax*, the entire insect stage occurs exclusively in the mouth parts from epimastigotes to metacyclics (Fig.1.3).

Table 1.1: Different trypanosome species and their corresponding hosts and diseases

Trypanosoma species	Host and disease where named
T. avium	Birds[a]
T. boissoni	Elasmobranchs e.g. sharks and rays[b]
T. brucei	Sleeping sickness in humans and nagana in cattle[c]
T. carassii	Freshwater fish (teleosts)[d]
T. cruzi	Chaga's disease in humans[c]
T. congolense	Nagana in cattle, horses, and camels[c]
T. equinum	Horses[e]
T. equiperdium	Dourine or covering sickness in horses[e]
T. evansi	Surra in cattle, camels and horses[f]
T. lewisi	Rats[g]
T. melophagium	Sheep[h]
T. percae	Fish[i]
T. rangeli	Human and cattle[j]
T. rotatorium	Amphibians[k]
T. simiae	Nagana in pigs, cattle, horses, camels and warthogs[l]
T. suis	Surra in cattle[f]
T. theileri	Ruminants[i]
T. triglae	Fish[b]
T. vivax	Nagana in cattle[c]

[a]Gutierrez (1989), [b]Jirku *et al.* (1995), [c]Barrett *et al.* (2003), [d]Overath *et al.* (1998), [e]Vickerman *et al.* (1993), [f]Hoare (1972), [g]Sturtevant and Balber (1983), [h]Büscher and Friedhoff (1984), [i]Gibson *et al.* (2005), [j]Grisard (2002), [k]Hysek and Zizka (1976), [i]Van Den Berghe and Zaghi (1963)

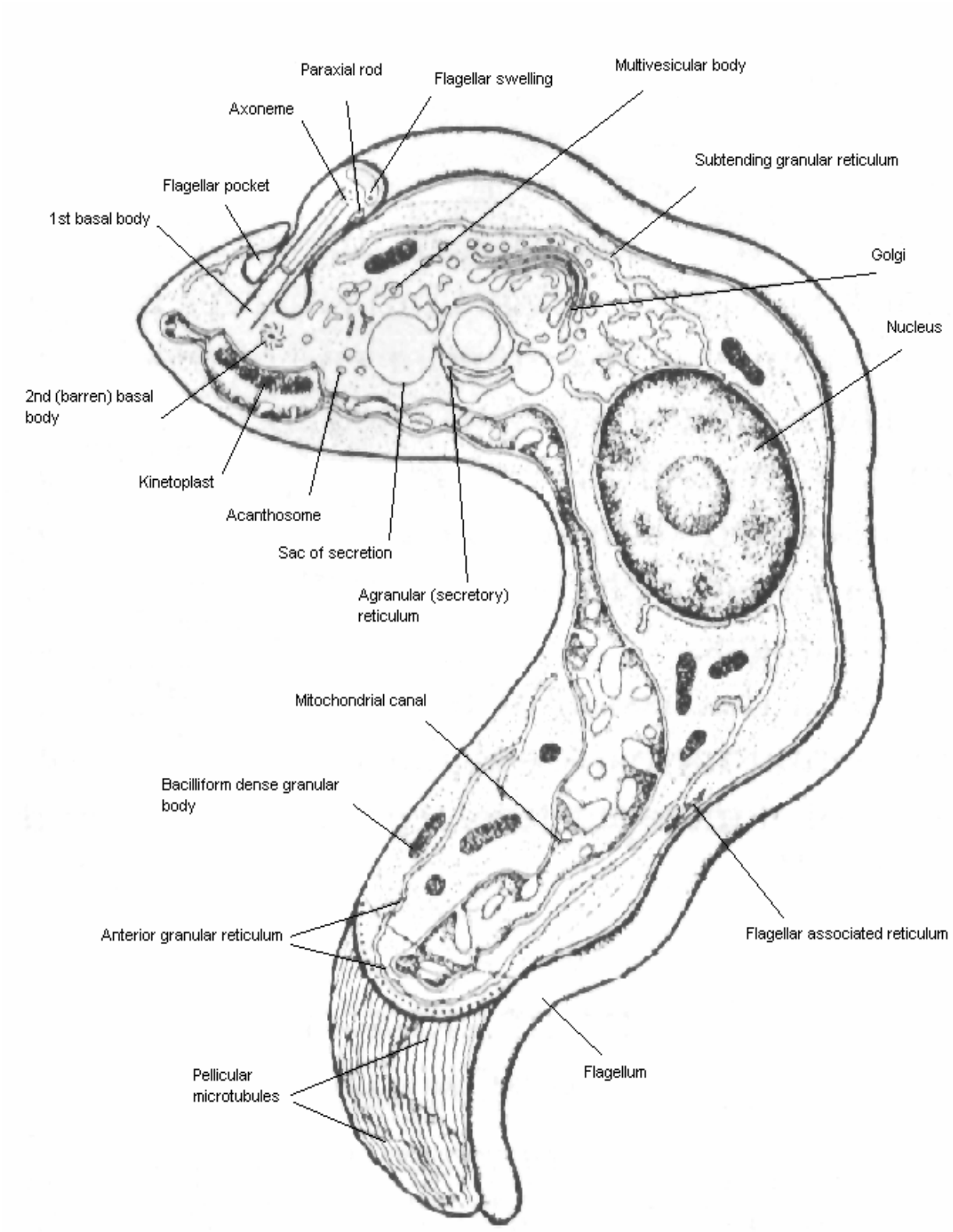


Figure 1.2 General structure of *T. congolense*. (Noble and Noble, 1982)

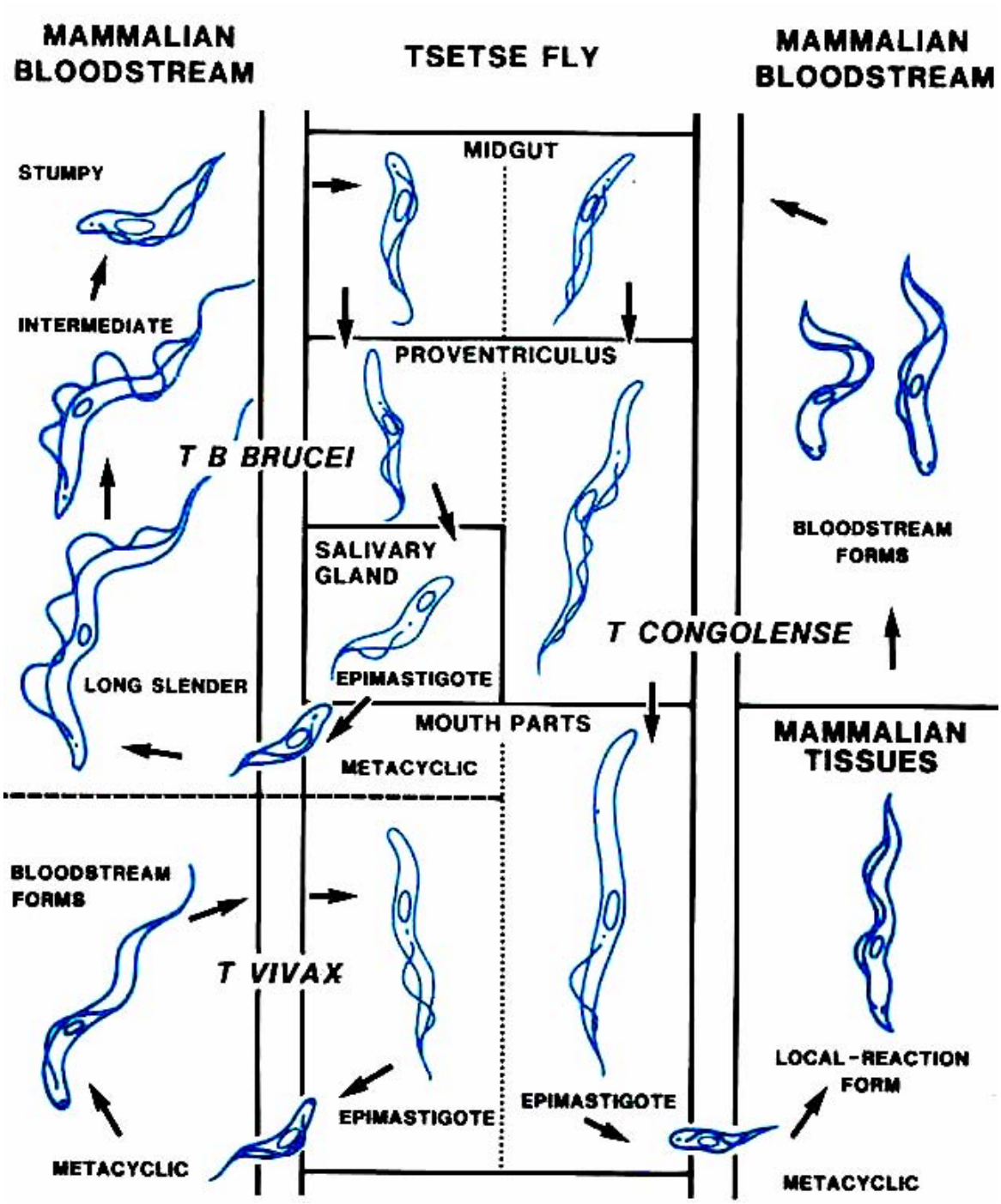


Figure 1.3 Life cycles of *T. b. brucei*, *T. congolense* and *T. vivax*. Parasites in mammalian bloodstream and tissues possess surface coats consisting of variable glycoprotein antigens, whereas those in insect stages have procyclin on the surface and are non-infective to mammals until the metacyclic stage found in the mouth parts of insects which are infective to mammals. *T. b. brucei* develops in the tsetse midgut, proventriculus and salivary glands, where metacyclic forms occur. *T. congolense* develops in the tsetse midgut, proventriculus and mouth parts, where metacyclic forms are produced. *T. vivax* develops only in the tsetse mouthparts. (http://www.ilri.org accessed on 15-02-2010).

1.2.3 Antigenic variation in trypanosomes.

During the bloodstream stage of the trypanosome life cycle, the parasite is covered with a monolayer of approximately 10^7 copies of a 60 kDa glycosyl-phosphatidyl inositol (GPI) anchored glycoprotein termed the variant surface glycoprotein (VSG) (Pays, 2006). VSG expression commences in the tsetse fly vector's salivary glands, at the metacyclic stage. This process is initiated just as division of parasites comes to an end resulting in metacyclics with a heterogeneous variable antigen type (VAT) referred to as metacyclic VAT (M-VAT) (Tetley *et al.*, 1987). Once in the mammalian host, the trypanosomes continue to express M-VAT and rapidly change from one VAT to another. Bloodstream form VATs are then expressed with one major VAT for each different peak of parasitaemia during infection (Hajduk and Vickerman, 1981). Only one predominant VSG antigen type is expressed on the surface of most parasites at any given time during an infection. VSG is a major parasite product that induces a strong immune response in the host which kills off most of the infecting parasites but a small number survive this challenge due to the different VSG they acquire using a process known as VSG switching (Pays *et al.*, 2004). This allows the parasites to escape antibody-mediated killing and multiply causing the next wave of parasitaemia and thus prolonging the infection. Antigenic variation is possible in infections initiated by a single trypanosome (Vickerman *et al.*, 1993).

The expression of different VSG types in a trypanosome population is controlled post-transcriptionally after polycistronic RNA is transcribed from a pool of approximately 1500 VSG genes and pseudo-genes (Taylor and Rudenko, 2006). Most of these genes are silent and only become active when they are linked to specialised VSG expression sites (VSG-ES). Expression sites are associated with several genes that are also expressed on the parasite surface such as the transferrin receptor and the serum resistance-associated gene in *T. b. rhodesiense* (Pays, 2006). VSG switching is facilitated by either *in situ* switching which shifts VSG-ES from one VSG gene to another, or through the homologous recombination between different VSG genes and pseudo-genes (Taylor and Rudenko, 2006). Most of the invariable stable antigens found in trypanosomes are not located on the parasite surface and consist of structural proteins and enzymes that are found at all stages of the trypanosome life cycle. Although they are not exposed to the host immune system, they can be highly immunogenic and important in immunopathology of the infection (Vickerman *et al.*, 1993) or for potential use in serodiagnosis. Examples include hsp-70 and *Tco*CATL (Authié *et al.*, 1993a). The procyclic form of *T. brucei* bears a dominant glycoprotein surface antigen called procyclin

which is synthesised and transported to the trypanosome surface before the VSG is lost during transformation of bloodstream forms to procyclics (Richardson *et al.*, 1988).

1.2.4 Immune mechanisms against trypanosomes and immunopathology

Mammalian immune reactions to trypanosome infection is mediated by both the low specificity pro-inflammatory innate response and the antigen specific adaptive immune response (Stijlemans *et al.*, 2007a). A tsetse bite transferring infective metacyclics into the host causes the classical activation of macrophages by trypanosomal DNA and the GPI anchor of the VSG leading to the release of pro-inflammatory cytokines like TNF, IL-12 and nitric oxide around the bite region and the formation of a chancre (Stijlemans *et al.*, 2007a; Baral, 2010). This results in an acute pro-inflammatory reaction by the host also known as a type I immune response which is responsible for the control of the first wave of parasitaemia. Because prolonged inflammation is pathological to the host, a type II anti-inflammatory immune response mediated by the production of type II cytokines such as IL-4, IL-10 and IL-13 follows the initial innate response (Baral, 2010). These molecules down-regulate classically activated macrophages and modulate them to become alternatively activated macrophages which are anti-inflammatory.

A strong proliferative response of B lymphocytes, first in the lymphoid system, and later in the central nervous system follows the initial chancre stage (Taylor, 1998). B-cell proliferation accounts for the destruction and clearance of parasites in successive waves of parasitaemia that follow the initial wave with high IgG and IgM titres observed that are involved in prolonging host survival during a chronic infection (Magez *et al.*, 2008). Along with the increase in B cell numbers is an accompanying decline in T cell populations and subsequent depression of the host's T cell-dependent responses (Taylor, 1998). Later during infection, as the disease becomes chronic, B lymphocytes become exhausted resulting in the absence of circulating IgG and reduced IgM levels leading to the suppression of host immune responses (Baral, 2010). Immunosuppression is responsible for the frequent incidence of concurrent viral, bacterial and parasitic infections found in trypanosome infected hosts (Askonas, 1985; Magez *et al.*, 2010). In addition, memory B-cell activation is permanently lost during infection when host B-cell compartments are irreversibly destroyed leading to failures in vaccination programmes against trypanosomosis and other diseases (Radwanska *et al.*, 2008; Magez *et al.*, 2010).

A notable pathological effect of trypanosome infections is anaemia (Taylor and Authié, 2004). Extravascular erythrocyte destruction is as a result of phagocytosis by the phagocytic mononuclear system which targets red blood cells with trypanosome antigen-antibody complexes on their surfaces. Phagocytosis is also thought to be mediated by trypanosome enzymes (Taylor and Authié, 2004). Immune complex formation and the release of active amines, peptides and lipids (collectively known as 'autocoids'), also play a major role in pathogenesis (Antoine-Moussiaux *et al.*, 2009). These are produced during the repeated destruction of trypanosome populations by host antibody, and by the shedding of antigens from trypanosomes into antibody-laden surroundings. The activation of the complement system by immune complexes on the surface of various cell types leads to the lysis or opsonisation of these cells and in the case of red blood cells leads to severe anaemia (Taylor, 1998). Autoantibodies against a wide variety of host tissue and cells is thought to be responsible for tissue lesions in the host, and the vast variety of VSG epitopes cross-react with host components leading to further tissue damage (Vickerman *et al.*, 1993).

1.2.4.1 Trypanotolerance

Trypanotolerance is the term used to describe animals that are resistant to the development of trypanosomosis after infection with trypanosomes (Murray *et al.*, 1982). Features describing trypanotolerance are the ability to control parasite proliferation, limit pathogenic effects of parasitaemia and the control of anaemia (Murray *et al.*, 2004). Trypanotolerance is common amongst game animals that in Africa act as reservoirs of the disease (Brown, 2008). Cape buffalo infected with trypanosomes develop lower levels of parasitaemia when compared to similarly-infected cattle with the action of xanthine oxidase and trypanosome specific IgG implicated in parasite control (Black *et al.*, 2001). West and Central African breeds of domestic cattle such as the N'Dama longhorn and West African shorthorn are resistant to trypanosome infection when compared to other breeds such as Zebu and European cattle (Tabel *et al.*, 2000). Resistant breeds of cattle are referred to as 'trypanotolerant, whereas breeds that succumb quickly to infection are referred to as 'trypanosusceptible'. Trypanotolerant cattle have lower parasitaemia and less anaemia than trypanosusceptible breeds (Murray *et al.*, 1982). Furthermore, there are differences in the immune response between the two breeds where studies have shown that N'Dama cattle have a greater ability to control the initial wave of parasitaemia than Zebu cattle and have more IgM secreting B-cells, proliferating $CD8^+$- and gamma-delta T-cells during a *T. congolense* infection (Taylor *et al.*, 1996; Tabel *et al.*, 2000). Studies have since shown that

the role of circulating IgM is limited in controlling successive waves of parasitaemia in a chronic infection and that trypanotolerance in cattle is associated with the secretion of IFNγ and TNF, the activation of macrophages and the production of nitric oxide; all hallmarks of an innate immune response during the initial phase of trypanosome infection (Magez *et al.*, 2008)

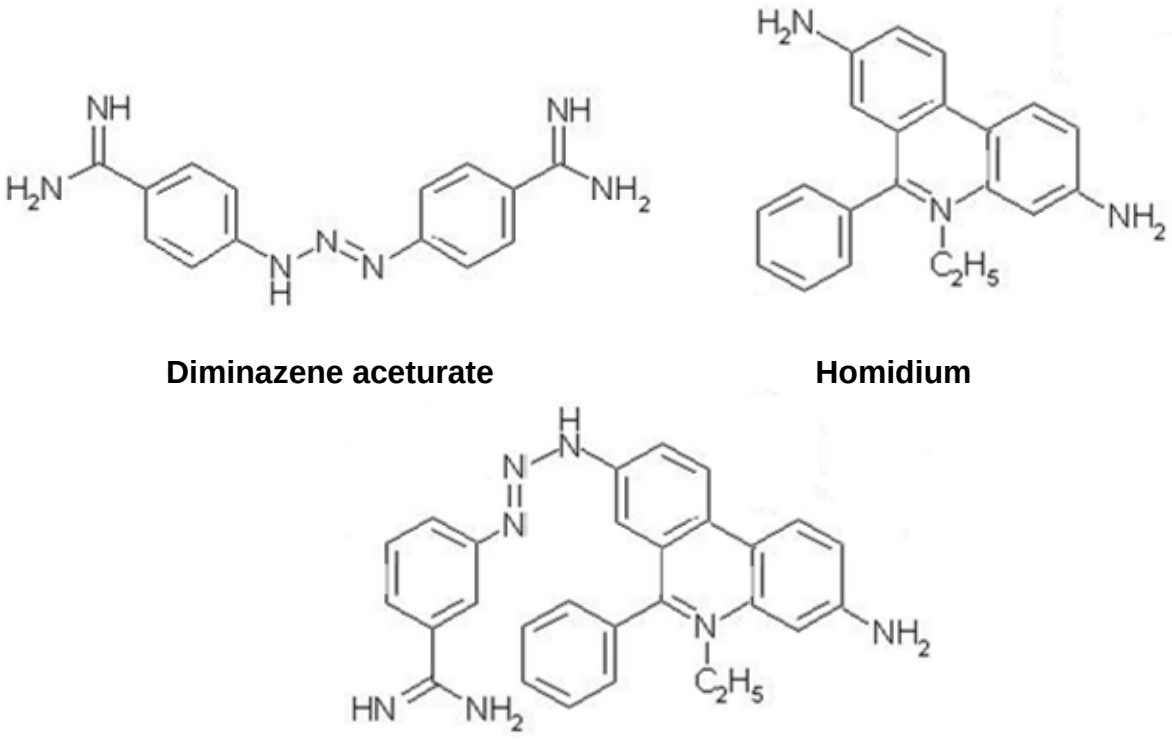
Experiments with bovine cross-breeds between resistant and susceptible cattle breeds showed that N'Dama and Boran cross-breeds were able to limit parasitaemia development (Naessens *et al.*, 2003). These results suggest that trypanotolerance arises as a result of two mechanisms: one which is an innate process independent of genetic origin that helps control parasite levels in the host and the other dependent on genetic origin and helps reduce anaemia and the death of white blood cells (Naessens *et al.*, 2003). Because trypanotolerant breeds of cattle have the ability to control parasitaemia and development of anaemia, their presence has been exploited in the humid and sub-humid zones of West and Central Africa. In areas of high trypanosome challenge, nearly all the cattle are of trypanotolerant breeds. Although smaller in size, trypanotolerant breeds are considered to be as productive as other breeds that are more susceptible to trypanosomes (Naessens, 2006).

1.3 CONTROL OF TRYPANOSOMOSIS

1.3.1 Chemotherapy

With the lack of surveillance and adequate veterinary structures in areas affected by trypanosomosis, disease control has been at best sporadic and inefficient. In sub Saharan Africa, individual farmers mainly use chemotherapy to manage the disease in livestock (McDermott and Coleman, 2001; Delespaux and de Koning, 2007). The use of drugs has been difficult to estimate, with many generic brands been sold unofficially rather than by monitored official routes (Geerts *et al.*, 2001). Three trypanocidal compounds have been used to treat veterinary trypanosomosis for over 50 years: diminazene aceturate, homidium and isometamidium. Isometamidium is synthesized by coupling homidium with a part of the diminazene molecule (Fig.1.4). Diminazene resistance is prevalent in areas where there is a high incidence of trypanosomosis and drug usage (Delespaux and de Koning, 2007). Resistance is associated with the P2 aminopurine transporter though other membrane transporter proteins are also implicated (Matovu *et al.*, 2003). Cross resistance between homidium and isometamidium has been reported (Delespaux and de Koning, 2007) and

occurs when large populations of infected cattle are repeatedly treated with isometamidium. The development of newer drugs that do not result in resistance include modifications of diminazene that contains furan linkers and aromatic diamidines (Delespaux and de Koning, 2007) and drugs that target unique glycosomal enzymes e.g. peroxins, not found in the mammalian host (Schliebs, 2006).



Diminazene aceturate **Homidium**

Isometamidium

Figure 1.4 Chemical structures of the three primary drugs used to treat African animal trypanosomosis. (Delespaux and de Koning, 2007).

1.3.2 Vector control

Early attempts targeting the tsetse fly vector used a combination of traps, bush clearance and arial spraying with dicholoro-diphenyl-trichloroethane (DDT) and benzene hexachloride (BHC) (Brown, 2008). This proved to be successful with the elimination of trypanosomosis in South Africa in 1953. Tsetse control using these methods is however not long-term as new cases of trypanosomosis have been identified in northern KwaZulu-Natal in the 1990s and as recently as 2009 (Kappmeier *et al.*, 1998; Mamabolo *et al.*, 2009). Methods for vector control such as aerial and ground spraying has been replaced by stationary odour-baited traps, pesticide

treated targets and cattle dip insecticides referred to as ‘pour-ons’ as a multifaceted approach to vector control (Brown, 2008).

Another suggestion for vector control is ‘transgenesis’ which is a method that reduces the ability of the insect to transmit trypanosomes and the development of vectors that interfere with the procyclic stage in the tsetse midgut (Aksoy, 2003; Magez *et al.*, 2010). Transgenesis also aims to reduce survival of the tsetse fly vector itself (Magez *et al.*, 2010). The development of transmission blocking vaccines (TBVs) is based on the idea that targeted antigens are less variable as they have not been under evolutionary pressure by the host immune system (Magez *et al.*, 2010). Tsetse flies fed on goats immunised with *in vitro* lysates from *T. brucei, T. congolense* and *T. vivax* were all able to control the development of respective parasites in their mid-guts (Murray *et al.*, 1985). A similar study reduced *T. b. rhodesiense* infection rates in flies fed on rabbits immunised with *G. pallidipes* midgut proteins (Kinyua *et al.*, 2005).

The dissemination of sterile males in tsetse infested areas have also been applied in controlling tsetse populations (Aksoy *et al.*, 2001; Hao *et al.*, 2001; Schofield and Maudlin, 2001; Aksoy, 2003). Using transgenic and sterile males is expensive and requires regular release of the flies. This method has proved to be successful in eradicating tsetse flies on the island of Zanzibar (Msangi *et al.*, 2000; Vreysen *et al.*, 2000). Long term eradication of the trypanosome vector has, however, been under evaluation as there is a possibility of re-infestation by new tsetse populations (Aksoy, 2003). Rolling out this scheme to mainland Africa or over a large land area greater than 10 million km^2 is not considered to be cost effective or sustainable (Rogers and Randolph, 2002).

1.3.3 Vaccine development

An area that has been explored for the control of trypanosomosis, but has borne numerous frustrations, is the development of vaccines (Radwanska *et al.*, 2008; Baral, 2010; Magez *et al.*, 2010). Cattle immunised with irradiated trypanosomes or purified VSG are immune to a homologous infection but are not protected from a heterologous challenge (Taylor and Mertens, 1999). Because trypanosomes have a very efficient method of switching GPI-anchored surface VSGs that are coded for by a wide array of genes and pseudo genes, it makes it difficult to use a VSG based vaccine against trypanosomosis (McCulloch and Horn, 2009). A more effective vaccine would be one that targets invariant antigens in the parasite

that do not change across infections (Authié *et al.*, 1993b; Taylor, 1998). Invariant antigens investigated as vaccine and chemotherapeutic targets include structural and surface parasite proteins and pathogenic trypanosome products released during infection.

1.3.3.1 Trypanosomal structural antigens

Trypanosomal structural proteins have been used as vaccine targets with varying success. Mice immunised with native tubulin from *T. brucei* conferred 60-80% protection against *T. brucei, T. congolense* and *T. rhodesiense* infections (Lubega *et al.*, 2002). This was confirmed in experiments using recombinant *T. evansi* beta-tubulin that partially protected against infections by *T. evansi, T. equiperdium* and *T. b. brucei* (Li *et al.*, 2007). Experiments using recombinant *T. brucei* actin showed similar degrees of protection against the same infecting strains as those used for recombinant beta-tubulin (Li *et al.*, 2009). Despite these promising results, further investigations are required to assess whether the immunity conferred by both beta-tubulin and actin is antigen-specific as a parallel control using other similarly recombinantly expressed trypanosomal proteins were not included in all three experiments (Magez *et al.*, 2010). Mice immunised subcutaneously with paraflagellar rod protein purified from *T. cruzi* epimastigotes were able to survive a challenge with 10^3 *T. cruzi* parasites (Wrightsman *et al.*, 1995). Interestingly, mice immunised intraperitoneally with purified paraflagellar rod protein succumbed to a challenge with an equal number of *T. cruzi* parasites (Wrightsman *et al.*, 1995). It was also noted that mice that received subcutaneous injections had a much lower anti-paraflagellar rod protein antibody titre when compared to those that were immunised intraperitoneally. These results suggested that protection was conferred by cell mediated, rather than antibody mediated immunity (Wrightsman *et al.*, 1995). Further experiments that measured the proliferation of splenocyte T cells and immunisation with recombinant paraflagellar rod protein revealed that protection is conferred by a Th-1-type cellular response (Miller *et al.*, 1996; Luhrs *et al.*, 2003).

1.3.3.2 Invariant surface antigens

Trypanosomal invariant surface antigens have also been assessed for suitability as vaccine targets. An initial study of the flagellar pocket membrane from different *T. b. rhodesiense* variable antigen types was able to localise flagellar pocket antigens on the surface of trypanosomes (Olenick *et al.*, 1988). This demonstrated that invariant surface proteins can be recognised by the infected host immune system. Immunisation of mice with preparations of

the flagellar pocket membranes conferred 60-40% protection for two of the three variable antigen types used for infection during the experiment (Olenick *et al.*, 1988). Further experiments using the flagellar pocket from *T. rhodesiense* as an antigen also provided partial protection to *T. congolense* and *T. vivax* infections in cattle (Mkunza *et al.*, 1995). Anti-flagellar pocket membrane antibodies raised in mice were used to screen a cDNA expression library of *T. brucei* bloodstream form parasites and predominantly identified heat shock protein 60 (HSP60) and the invariant surface glycoprotein 75 (ISG75) (Radwanska *et al.*, 2000a). Mice infected with *T. brucei* revealed that an anti-HSP60 response is induced when parasites switched into stumpy forms, whereas a weak antibody response was detected for ISG70 (Radwanska *et al.*, 2000b).

Based on these observations, ISG75 recombinantly expressed in *E. coli* was used to immunise rabbits and the serum collected was able to detect all of the different species in the *Trypanozoon* subgenus, proving that ISG75 is conserved across the genus (Tran *et al.*, 2008). However, ISG75 immunised mice were unable to survive an infection with *T. brucei* parasites (Magez *et al.*, 2010). In experiments using mice immunised with the flagellar pocket membrane, partial protection was only observed in mice challenged with less than 10^3 parasites. An infection with more than 10^3 parasites suppressed specific antibody responses to ISG70 with antibody titres dropping rapidly after two weeks of infection (Radwanska *et al.*, 2000b). Furthermore, although immunisation with invariant surface antigens resulted in the production of antigen specific memory B-cells, a subsequent infection with trypanosomes seemed to abolish this memory in addition to other B-cell compartments leaving the host susceptible not only to trypanosomosis but also to other livestock diseases such as foot and mouth disease, anthrax and swine fever (Radwanska *et al.*, 2008; Magez *et al.*, 2010).

1.3.3.3 Pathogenic products from trypanosomes

Because of antigenic variation, the host immune system is constantly producing antibodies to different antigen types in one infection clearing most parasites but in the process releasing many trypanosome products into the host circulatory system (Antoine-Moussiaux *et al.*, 2009). These products include lipids degraded by trypanosome phoslipase A_1 generating free fatty acids that disrupt cell membranes (Wainszelbaum *et al.*, 2001), hemolysins that assist *T. cruzi* invade cells (Andrews and Whitlow, 1989) and mitogens that enhance non-specific polyclonal B cell proliferation enabling parasites evade a specific immune response (Reina-San-Martin *et al.*, 2000). Other factors are also responsible for inflammation and toxicity

include peptidases such as CATLs and oligopeptidases (Authié *et al.*, 1993a; Coetzer *et al.*, 2008). This has led to exploring the development of vaccines that target products released during an infection in a strategy referred to as an 'anti-disease' approach (Authié *et al.*, 2001). A trypanosomal anti-disease vaccine is formulated to alleviate the symptoms associated with the disease rather than attempting to eradicate the infecting parasite itself.

Early during infection as parasites divide in the host, trypanosomes rapidly switch their VSG coat and in the process release large amounts of GPI-anchored VSG, through the action of trypanosomal phospholipase C (Pays *et al.*, 2004). The glycosyl-inositol-phosphate (GIP) moiety of the GPI anchor which is attached to soluble VSG is in itself a strong activator of macrophages that produce TNF (Magez *et al.*, 2002). TNF is responsible for the pro-inflammatory response observed in a type I immune response during infection and excessive amounts circulating in the host lead to TNF-mediated trypanosomosis associated immunopathology (Magez *et al.*, 2002). Immunisation of mice with native GPI from *T. brucei* was able to reduce pathology associated with trypanosome infection in mice by reducing TNF-associated symptoms, reduce anaemia and increase the lifespan of immunised mice by approximately 10 days when compared to non-immunised mice (Stijlemans *et al.*, 2007b). Because phospholipase C is responsible for the shedding of parasite VSG, mice infected with phospholipase C null mutant parasites were more resistant to disease, had lower parasitaemia and increased survival when compared to wild type infections (Baetselier *et al.*, 2001). Disease resistance in null mutant infections is attributed to mice been able to switch from a classical type I macrophage activation reaction to the alternative type II reaction that does not involve overproduction of TNF observed in wild type infected mice (Namangala *et al.*, 2001). Phospholipase A_1 in *T. brucei* and *T. cruzi* infections is responsible for the breakdown of phospholipids into free fatty acids and lysophospholipids (Opperdoes and van Roy, 1982; Wainszelbaum *et al.*, 2001). The release of free fatty acids and lysophospholipids leads to the degradation of phospholipid bilayers in surrounding cells and cause inflammation that is associated with lesions observed in trypanosomosis.

During *T. congolense, T. vivax* and *T. evansi* infections, sialidases have been implicated in causing anaemia (Nok and Balogun, 2003; Nok *et al.*, 2003; Antoine-Moussiaux *et al.*, 2009). Sialidases are responsible for the removal of sialic acid from cell surfaces and in *T. congolense* infections, an increase in parasitaemia is associated with an increase in serum free sialic acid that is released from the surface of erythrocytes (Murray *et al.*, 1985). Red

blood cells without sialic acid on their surfaces are targeted for erythrophagocytosis by macrophages leading to late stage anaemia in *T. congolense* infections (Antoine-Moussiaux *et al.*, 2009). The down-regulation of sialidases in *T. congolense* resulted in reduced rates of infectivity with a longer pre-patent periods, lower parasitaemia and less disease associated pathogenesis in mice infected with parasites targeted for sialidase RNAi *in vivo* and knock-out (Coustou *et al.*, 2011). Mice immunised with recombinant sialidases from *T. congolense* are also able to survive longer than non-immunised mice when challenged with either *T. congolense* IL1180 or IL3000 (Coustou *et al.*, 2011).

Experiments using trypanotolerant and trypanosusceptible cattle challenged with *T. congolense* revealed a 69 kDa antigen which elicited an IgM response from both breeds of cattle and an IgG response from trypanotolerant cattle against a 33 kDa antigen (Authié *et al.*, 1993a; Authié *et al.*, 1993b). This 33 kDa antigen was identified as a cysteine peptidase referred to as *Tco*CATL (Authié *et al.*, 1992). Cattle immunised with *Tco*CATL recovered faster and showed a higher IgG response to *T. congolense* challenge than non-immunised cattle. Another peptidase that has been implicated in host pathology is the serine peptidase oligopeptidase B (OPB) which is released by dead and dying parasites and found to be active in the plasma of rats with an acute infection of *T. b. brucei* (Morty *et al.*, 2001). The role that these two peptidases play during trypanosome infections is discussed further in Section 1.4

1.4 PROTEOLYTIC ENZYMES IN TRYPANOSOMES

Enzymes that catalyse the hydrolysis of peptide bonds are collectively referred to as proteolytic enzymes (proteases) or peptidases (Barrett and Rawlings, 2004; Rawlings *et al.*, 2010). With the discovery of a seventh catalytic type of proteolytic enzyme, the asparagine peptide lyases (Rawlings *et al.*, 2011), the terms peptidase and proteolytic enzyme are no longer synonymous because they do not use hydrolysis to cleave peptide bonds. These self-cleaving enzymes use asparagine as the nucleophile in the lyase reaction (Rawlings *et al.*, 2011). Peptidases are further classified as exopeptidases that cleave proteins at the N- or C-termini (aminopeptidases and carboxypeptidases respectively) or endopeptidases that hydrolyse internal peptide bonds (Barrett and Rawlings, 2004). Endopeptidases have also been referred to as proteinases (Barrett and McDonald, 1986). Seven classes of proteolytic enzymes are distinguished according to their respective reactive groups that catalyse the cleavage of peptide bonds (Table 1.2). The first letter of the reactive group is used for the

name of proteolytic enzyme families and clans (Rawlings *et al.*, 2008). A family is a group of proteolytic enzymes that share significant amino acid identity to a representative proteolytic enzyme e.g. prolyl oligopeptidase in the case of the prolyl oligopeptidase family of serine peptidases, (Barrett and Rawlings, 2004; Rawlings *et al.*, 2008). Members of the same family possess the same catalytic group as the representative proteolytic enzyme. A clan on the other hand, can have more than one family and members have similar tertiary structures and arrangement of reactive groups (Barrett and Rawlings, 2004).

Table 1.2: The seven different classes of proteolytic enzymes (Rawlings *et al.*, 2010).

Class	Reactive group	Example	Inhibitors
Serine	Hydroxyl group of serine	Oligopeptidase B[a,b,c]	3,4-dichloroisocoumarin[a,b,c] peptidyl-chloromethylketones[d], protamine[e]
Threonine	Hydroxyl group of threonine	Macropain[f] (20S Proteasome)	Epoxomicin[f]
Cysteine	Thiol group of Cysteine	Papain, falcipain[g], *Tco*CATL[j] Metacaspases[k]	Cystatins[h,i] stefins[h,i], Cysteine peptidase inhibitors (ICPs)[l]
Aspartic acid	Carboxyl group of aspartic acid	Pepsin[m]	Pepstatin A[m]
Metallo	Zn^{2+}/ Ca^{2+}	Matrix metallopeptidases (MMPs)[n], *E. coli* aminopeptidase N[o] *T. congolense* aminopeptidases[p] *T. cruzi* aminopeptidase[q]	Bestatin[r] Doxycycline[s] tissue inhibitors of metalloproteinases (TIMPs)[t]
Glutamic acid	Carboxyl group of glutamic acid	*T. emersonii* glutamic peptidase 1[u]	Transition state analog inhibitor (TA1)[v]
Asparagine peptide lyases	Asparagine	*E. coli* Tsh autotransporter[w]	Undetermined[w]

[a]Morty *et al.* (2005a), [b]Coetzer *et al*. (2008), [c]Morty and Burleigh (2004), [d]Morty *et al*. (2000), [e]Tsuji *et al*. (2006), [f]Mordmüller *et al.* (2006), [g]Kumar *et al.* (2007), [h]Turk *et al*. (2008), [i]Abrahamson *et al*. (2003), [j]Authié *et al*. (1992), [k]Proto *et al.* (2011), [l]Sanderson *et al.* (2003), [m]Yoshida *et al.* (2006), [n]Solberg *et al*. (2003), [o]Ito *et al* (2006), [p]Pillay (2011), [q]Cadavid-Restrepo *et al*. (2011), [r]Ashmun and Look (1990), [s]Stechmiller *et al.* (2010), [t]Brew *et al*. (2000), [u]O'Donoghue *et al*. (2008), [v]Pillai *et al.* (2007), [w]Rawlings *et al*. (2011).

Proteolytic enzymes show a preference for particular amino acid residues adjacent to the peptide bond they cleave (Berger and Schechter, 1970; Barrett and Rawlings, 2004). The substrate residue immediately or terminal to the scissile bond is called P_1 with subsequent residues labelled P_2, P_3….P_n towards the N-terminus and $P_{2'}$, $P_{3'}$….$P_{n'}$ towards the C-terminus of the scissile bond. The corresponding sub-sites on the proteolytic enzyme are labelled S1 with subsequent sub-sites towards the N-terminus of the proteolytic enzyme labelled as S_2, S_3….S_n (Fig. 1.5). Sub-sites binding to the prime residues of the substrate are labelled $S_{1'}$, $S_{2'}$…$S_{n'}$.

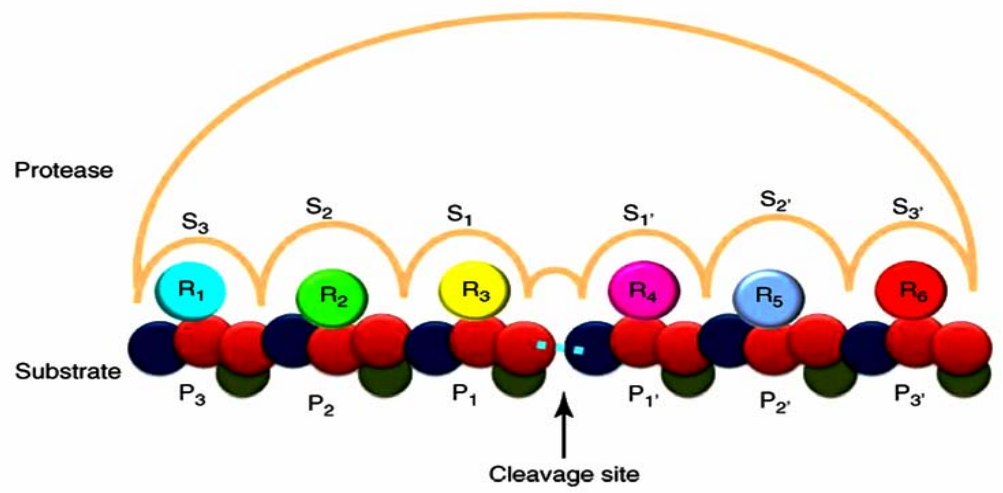


Figure 1.5 Berger and Schechter model for substrate specificity. The scissile bond is the site of cleavage with S_3 - $S_{3'}$ the peptidase sub-sites and P_3 - $P_{3'}$ the corresponding substrate amino acid residues with R_1-R_6 side chain groups that fit into the proteolytic enzyme sub-sites. (Smooker *et al.*, 2010)

1.4.1 Cysteine peptidases

Cysteine peptidases possess a cysteine residue in their active site. There are 35 known families in this class of peptidase (Rawlings *et al.*, 1994; Rawlings *et al.*, 2008), all with cysteine and histidine residues in their catalytic site. The order of cysteine and histidine residues in the primary structure differs between different families (Barrett and Rawlings, 2004). Cysteine peptidases are divided into several clans. Clans CD and CC are comprised of viral peptidases and clan CB contains legumain-like cysteine peptidases e.g. *T. brucei* GPI-transamidase (Nagamune *et al.*, 2003). Clan CA is the largest group of cysteine peptidases containing the papain family C1, with papain as the representative peptidase and family C2 of calpain-like cysteine peptidases which possess a calcium binding regulatory domain (Sajid and McKerrow, 2002). The papain-like family C1 is further divided into two major sub-families: cathepsin-B like peptidases which possess an occluding loop (Mendoza-

Palomares *et al.*, 2008) and cathepsin-L like peptidases which bear a unique 'ERFNIN' amino acid residue sequence (Sajid and McKerrow, 2002).

T. congolense cysteine peptidase, *Tco*CATL, is a member of the cathepsin-L like cysteine peptidase family (Lecaille *et al.*, 2002). Papain-like cysteine peptidases possess a highly conserved catalytic site formed by three catalytic residues: Cys^{25}, His^{159} and Asn^{175} (papain numbering) (Rawlings *et al.*, 2008; Caffrey and Steverding, 2009). Cys^{25} and His^{159} form an ion pair stabilised by Asn^{175} via a hydrogen bond. During peptide hydrolysis, the nucleophilic cysteine attacks the carbonyl carbon of the scissile bond forming a tetrahedral intermediate (Fig 1.6). The tetrahedral intermediate transforms into an acyl enzyme releasing the C-terminal portion of the peptide. The acyl enzyme is hydrolysed by water and forms another tetrahedral intermediate that dissociates into the free enzyme and releases the N- terminal portion of the peptide (Lecaille *et al.*, 2002; Barrett and Rawlings, 2004) .

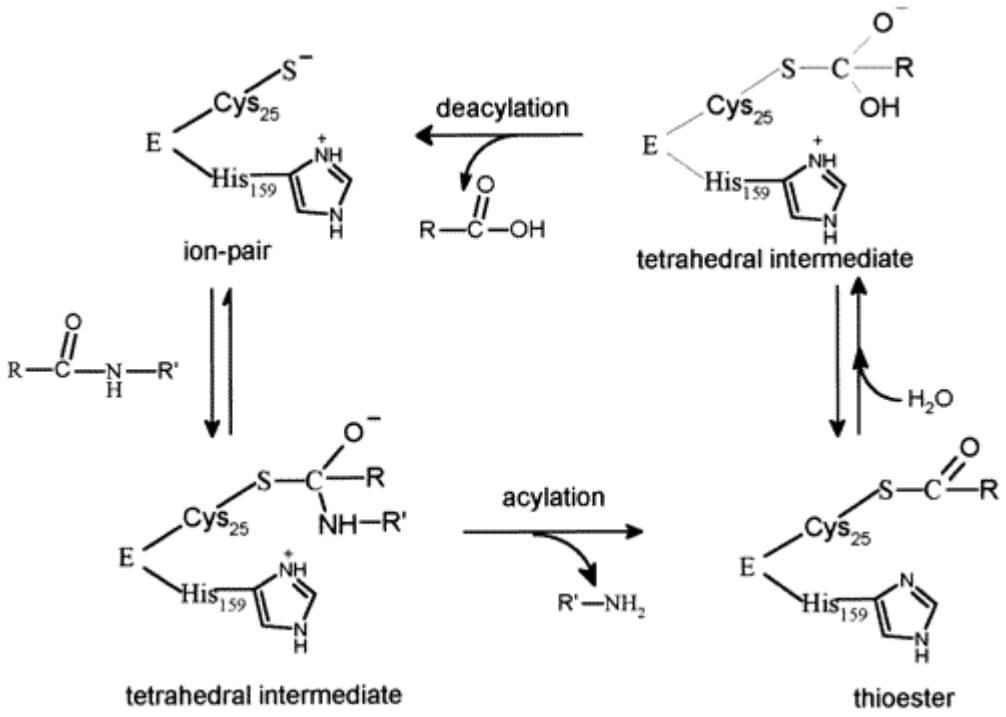


Figure 1.6 Proposed reaction mechanism for cysteine peptidases. (Lecaille *et al.*, 2002).

1.4.1.1 The major cysteine peptidase in *T. congolense* (*Tco*CATL)

So far, two closely related cysteine peptidase families have been identified in *T. congolense*, CP1 (Fish *et al.*, 1995) and CP2 (Jaye *et al.*, 1994) with both families sharing 90% sequence identity (Boulangé *et al.*, 2001). Congopain or *Tco*CATL, using a more recent nomenclature proposed by Caffery and Steverding (2009), is a 33 kDa CP2- type lysosomal peptidase that

has been isolated from *T. congolense* bloodstream forms (Authié *et al.*, 1992; Mbawa *et al.*, 1992). In contrast to its *T. cruzi* homologue, cruzipain (*Tcr*CATL), that retains activity below pH 4 (Chagas *et al.*, 1997), *Tco*CATL shows maximum activity at pH 6.4 which rapidly decreases at a lower pH. This difference in activity is attributed to the presence of Leu (*Tco*CATL) in place of Glu (*Tcr*CATL) at position 205 (papain numbering) at the S2 subsite (Chagas *et al.*, 1997). *Tco*CATL is expressed as a 444 amino acid pre-pro-peptide composed of a signal peptide (20 amino acids long), a pro-peptide region (105 amino acids long containing a 56 amino acid inhibitor domain), and a catalytic domain that is 319 amino acid residues long (Fig.1.7) (Lecaille *et al.*, 2002). The pro-peptide, that contains a 56 amino acid residue long inhibitory domain, is necessary for the production of mature *Tco*CATL as it blocks the active site during protein folding to prevent proteolytic activity (Lalmanach *et al.*, 1998). Also present is a 130 amino acid long C-terminal extension linked to the catalytic domain by a polyproline hinge, a feature also present in cruzipain (Cazzulo and Frasch, 1992; Boulangé *et al.*, 2001) but absent in mammalian cathepsins. The C-terminal region is thought to be highly immunogenic and is similar to the C-terminal domain found in cruzipain which is released from the rest of the mature protein during infection (Martinez *et al.*, 1991). Due to its catalytic action, mature *Tco*CATL is found in the lysosomes of living parasites and is only released after the host immune system lyses the parasite (Mbawa *et al.*, 1991b).

*Tco*CATL is a highly immunogenic antigen of *T. congolense* and induces both humoral and cellular responses in cattle (Authié *et al.*, 1993a). Because of its immunogenicity, *Tco*CATL has been identified as a pathogenic factor during trypanosomosis (Authié *et al.*, 2001). Truncated CP1 and CP2 lacking the C-terminus (due to its highly immunogenic nature) were used to immunise cattle in order to assess whether antibodies against the catalytic domain could inhibit trypanosome CPs during infection (Authié *et al.*, 2001). Immunised cattle subsequently challenged with *T. congolense* were assessed for anaemia and antibody responses. Although the onset of anaemia was the same in all infected animals, immunised cattle were able to recover from anaemia and regain normal leukocyte counts 2-3 months after challenge. Immunised cattle did not suffer from severe weight loss observed in non-immunised controls (Authié *et al.*, 2001). This demonstrated that both CP1 and CP2 have a role to play in anaemia and more so for CP2 in immunosuppression. Immunised cattle were also able to mount a high IgG response against native CP and VSG antigens during infection, which seemed to protect them from pathology associated with the disease (Lalmanach *et al.*, 2002). It has also been shown that anti-CP antibody titres are more specific to native CP

when *Tco*CATL is complexed with α_2-macroglobulin as an adjuvant (Huson *et al.*, 2009). This is because *Tco*CATL retains its 3-dimensional conformation inside the α_2-macroglobulin cage-like structure thus increasing the efficiency of delivery to antigen presenting cells (Huson *et al.*, 2009). *Tco*CATL has some unique characteristics when compared to other trypanosomal cathepsin-L like cysteine peptidases. Recombinant *Tco*CATL expressed in *P. pastoris* forms a dimer at neutral pH but remains as a monomer at acidic pH (Boulangé *et al.*, 2011). Also, several other *Tco*CATL-like peptidases with non-classical active site motifs have been identified by comparative genomics and were shown to be active at physiological pH (Pillay *et al.*, 2010). Both these factors are important when considering the development of an anti-disease vaccine as cognisance needs to be taken of dimer-specific epitopes that are recognised by trypanotolerant cattle and the activity of *Tco*CATL-like peptidases that may not be inhibited in cattle only immunised against classical *Tco*CATL.

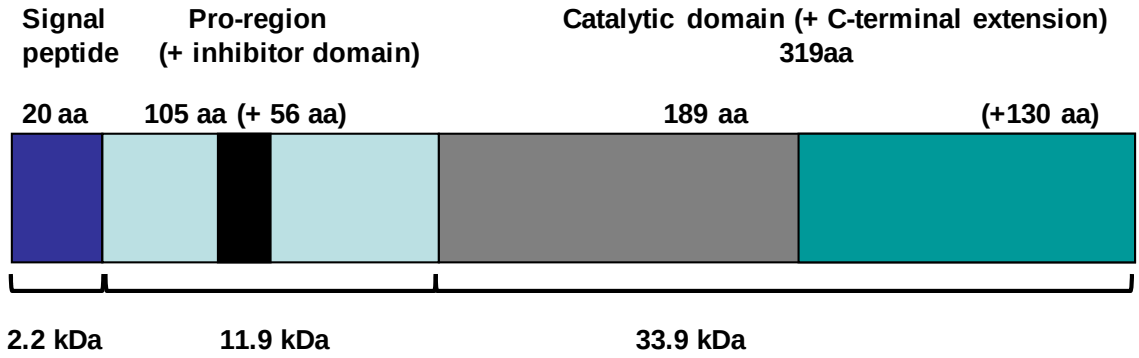


Figure 1.7 Structural domains of *Tco*CATL. *Tco*CATL is expressed as a 48 kDa, 444 amino acid long pre-pro-enzyme with a signal peptide, a pro-region containing an inhibitor sequence and a catalytic domain with a C-terminal extension linked by a polyproline hinge.

1.4.2 Serine Peptidases

Serine peptidases are a class of peptidases that possess a serine residue in their active site. Over a third of all known proteolytic enzymes are serine peptidases which have been identified across different species and are divided into 50 clans and 184 families (Rawlings *et al.*, 2008). The clans are grouped according to the protein fold surrounding the catalytic site e.g. clan PA peptidases possess Greek-key β-barrel protein folds, clan SB consists of a 3-layer sandwich fold and clan SC an α/β hydrolase fold around the active site (Page and Di Cera, 2008). In many serine peptidases, the catalytic action of the serine hydroxyl group is improved by the presence of a histidine residue and a stabilising aspartate residue, with the side chains of serine, histidine and aspartate building the catalytic triad (Barrett and Rawlings, 2004). During hydrolysis, the nucleophilic hydroxyl group of serine attacks the carbonyl carbon of the peptide bond under hydrolysis and forms a tetrahedral intermediate that

dissociates into an acyl-enzyme intermediate and releases the N-terminal portion of the cleaved peptide (product 1, Fig. 1.8). Addition of water generates a second tetrahedral intermediate that regenerates the original catalytic triad releasing the C- terminal portion of the cleaved peptide (product 2, Fig. 1.8) (Rea and Fülöp, 2006). Members of clan SC are α/β hydrolase-fold enzymes consisting of parallel β-strands surrounded by α-helices (Page and Di Cera, 2008). Both endo- and exo- peptidases are represented in this clan such as members of family S33 that hydrolyse N-terminal proline (exo-peptidase) and those in family S9 that cleave peptide bonds within peptides (endo-peptidase) (Rea and Fülöp, 2006). Family S9 is represented by prolyl oligopeptidase (EC 3.4.21.26) with various other members including dipeptidyl peptidase IV (EC 3.4.14.5), acylaminoacid releasing enzyme (EC 3.4.19.1) and oligopeptidase B (EC 3.4.21.83) (Barrett and Rawlings, 2004; Rawlings *et al.*, 2008).

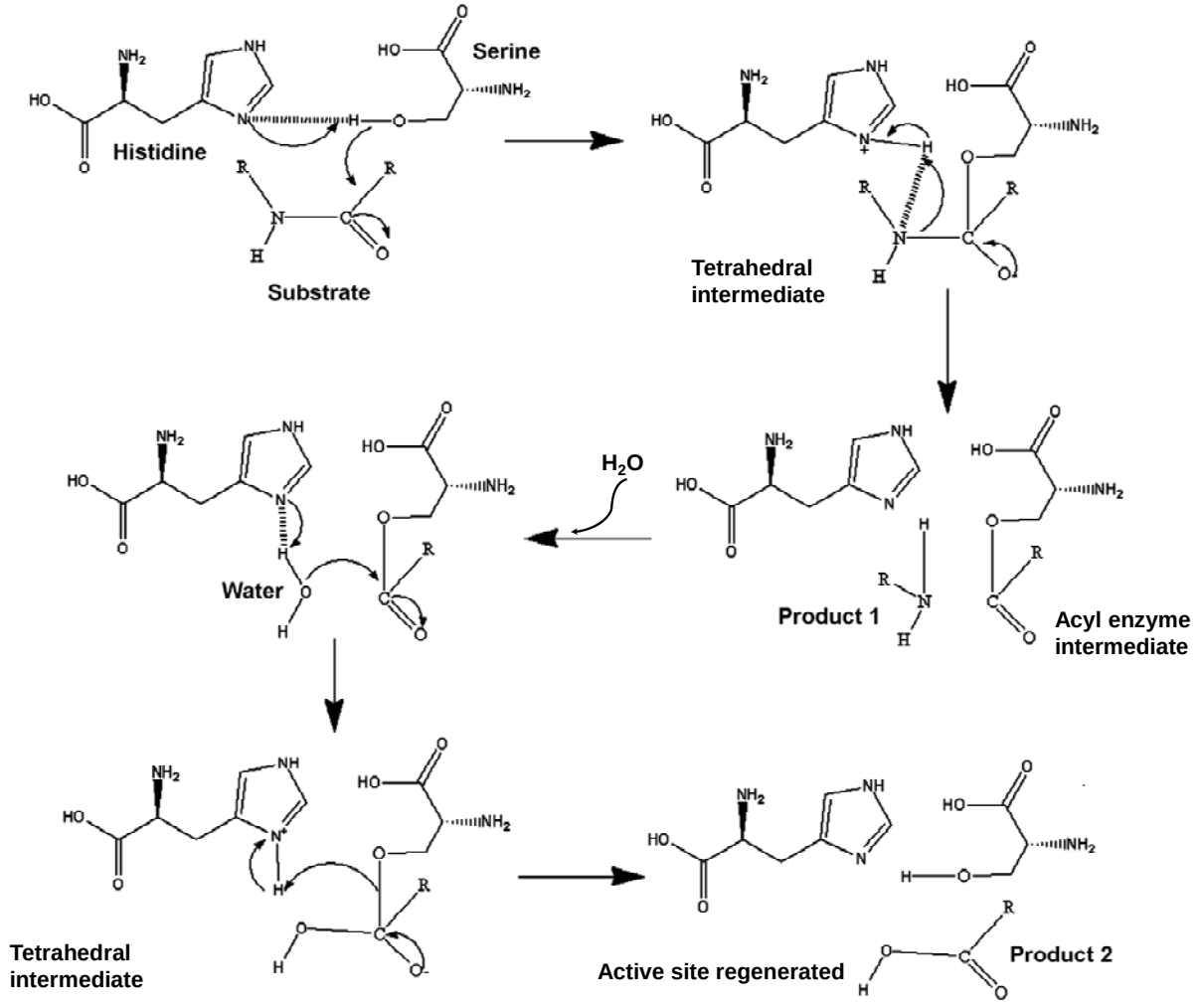


Figure 1.8 Mechanism of action for serine peptidases. (Voet *et al.*, 2006).

1.4.2.1 Oligopeptidase B

Oligopeptidase B (OPB) is a serine peptidase from the prolyl oligopeptidase family that hydrolyses peptide bonds of low molecular mass peptides smaller than 30 amino acids long (Barrett and Rawlings, 2004). The representative peptidase, prolyl oligopeptidase, cleaves peptides on the carboxy end of proline residues, whereas OPB only hydrolyses peptides on the carboxy end of basic residues, preferring arginine over lysine (Kanatani *et al.*, 1991; Morty *et al.*, 1999a; Coetzer *et al.*, 2008). The structure of oligopeptidase B has been described in *E. coli* (Gerczei *et al.*, 2000) through homology modelling using prolyl oligopeptidase as a template. The crystal structure of OPB from *L. major* (*Lm*OPB) has also recently been described in complex with antipain, an OPB inhibitor (McLuskey *et al.*, 2010). *Lm*OPB is a two-domain structure with a C-terminal catalytic domain and an N-terminal *β*-propeller domain linked by a hinge region (Fig. 1.9) (McLuskey *et al.*, 2010). The C-terminal catalytic domain contains the catalytic triad Ser^{577}, Asp^{662} and His^{697} located on loops between *β*-strands facing the propeller creating a cavity where antipain is bound (McLuskey *et al.*, 2010). The N-terminal regulatory domain is made up of a seven bladed *β*-propeller that lies over the catalytic triad allowing only short peptides access to the active site and consequently endogenous serine peptidase inhibitors (serpins) do not inhibit prolyl oligopeptidases (Fülöp *et al.*, 1998)

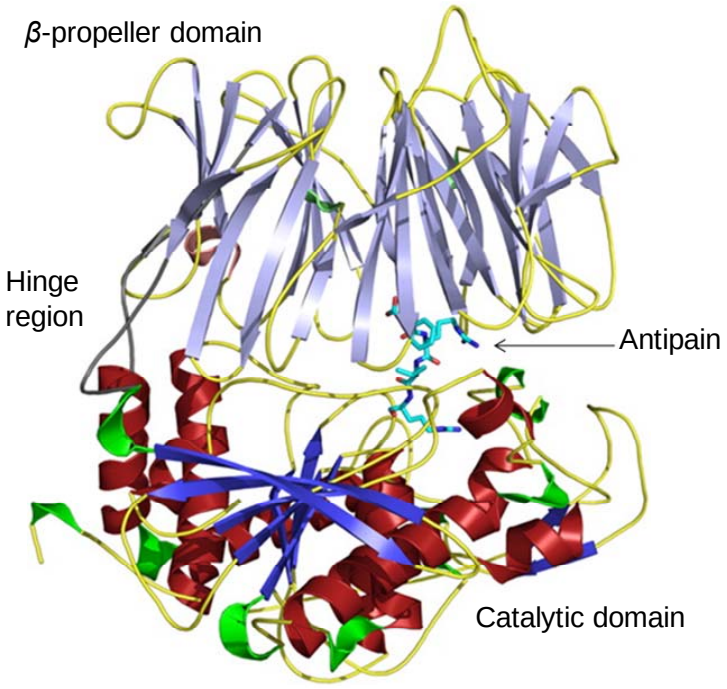


Figure 1.9 The structure of *Lm*OPB crystallised with antipain. General loop regions are shown in yellow, α-helices in red, β-strands in blue, and 3_{10} helices in green. ← shows the position of antipain in cyan as a stick model bound to the catalytic domain. [Adapted from McLuskey *et al.* (2010)].

Trypanosomal oligopeptidase B efficiently hydrolyses peptides that contain basic amino acid residues (Arg or Lys) in P_1 with a preference for arginine residues in both P_1 and P_2, favouring basic residues over Gly, Phe, Leu, Thr and Pro (Morty *et al.*, 1999a; Morty *et al.*, 1999b; Coetzer *et al.*, 2008). Natural peptides such as host peptide hormones and peptides with a high number of basic residues are efficiently hydrolysed by OPB (Troeberg *et al.*, 1996b; Morty *et al.*, 2005a). *In vitro*, OPB hydrolyses the peptide hormone atrial natriuretic factor (ANF), at four sites, thereby reducing ANF smooth muscle relaxant and hypertensive properties (Morty *et al.*, 2005a). Other hormones hydrolysed by OPB include adrenocorticotropic hormone (ACTH) (Tsuji *et al.*, 2004), neurotensin and reduced vasopressin (Troeberg *et al.*, 1996a). Oligopeptidase B is found in the cytoplasm and has been isolated from several species of trypanosomes, i.e. *T. b. brucei* (Troeberg *et al.*, 1996a), *T. congolense* (Morty *et al.*, 1999a), *T.* evansi (Morty *et al.*, 2005a) and *T. vivax* (Huson, 2006). Oligopeptidase B is released by dead or dying parasites into the host circulation where it retains full catalytic activity (Morty *et al.*, 2001). This is significant in host pathology since the degradation of peptide hormones could damage physiological homeostasis and subsequently cause several symptoms associated with infection (Troeberg *et al.*, 1996a; Morty *et al.*, 2005a). The trypanocidal agents suramin, diminazene and pentamidine inhibit OPB activity (Morty *et al.*, 1998; Coetzer *et al.*, 2008). Pentamidine inhibited OPB in a competitive manner whereas suramin inhibition was non-competitive in a fluorescent peptide substrate assay using native OPB. Inhibition of OPB by suramin and its analogues showed a relationship with trypanosome death in culture. This strongly suggests that OPB is one of the targets of suramin activity in trypanosomes (Morty *et al.*, 2005c). Inhibiting OPB with peptidyl-chloromethylketones and peptidyl-phosphonate diphenyl esters killed *in-vitro* cultures of *T. b. brucei*. A peptidyl-phosphonate diphenyl ester also cured mice after infection with trypanosomes (Morty *et al.*, 2000). In *T. cruzi*, oligopeptidase B has been implicated in the invasion of mammalian cells during infection (Burleigh *et al.*, 1997; Caler *et al.*, 1998). It was found that *T. cruzi* OPB was involved in the recruitment of host lysosomes by Ca^{2+}-signalling activity (Burleigh *et al.*, 1997). Deletion of the gene coding for OPB in *T. cruzi* affected the ability of null-mutant *T. cruzi* parasites to invade host cells *in vitro* and establish an infection in mice (Caler *et al.*, 1998). The addition of recombinant OPB to null mutant parasite extracts restored Ca^{2+}-signalling activity, proving that OPB has a central role to play in *T. cruzi* infections (Caler *et al.*, 1998).

Gene knock-out studies in *Leishmania major* and *L. donovani* showed that the loss of OPB resulted in delayed foot pad lesion formation in mice but no difference was observed in virulence during infection (Munday *et al.*, 2011; Swenerton *et al.*, 2011). Knock-out parasites were also less able to infect and survive within macrophages *in vitro* and a prominent up-regulation of enzymatically inactive membrane associated enolase was detected in *L. donovani* parasites without OPB (Swenerton *et al.*, 2011). Macrophages infected with OPB knock-out parasites when compared to wild type parasites showed significant up-regulation of different families of proteins including those involved in cytokine secretion, signal transduction, and the inflammatory response. The absence of enolase on *Leishmania* membrane surfaces enables the parasite to bypass macrophage activation during early stages of infection and knock-out parasites proved that OPB has a role to play in clearing membrane bound enolase (Swenerton *et al.*, 2011).

To study the roles that trypanosome peptidases play in parasite survival and infection, it is necessary to employ techniques that either down-regulate or completely remove their occurrence in the parasite. Down-regulation of a gene can be achieved at the transcription level by targeting RNA which is used as a template during translation whereas deletion of the gene is carried out at DNA level. These techniques used to identify gene function fall under the broad area of reverse genetics.

1.5 ELUCIDATION OF GENE FUNCTION IN PROTOZOAN PARASITES USING REVERSE GENETICS

Traditional genetics (also called forward genetics) has relied on the identification of genes based on a phenotype that has been previously described e.g. as a disease or mutation. Following advances in gene sequencing, the sequences of several genes have been elucidated but their functions are yet to be understood (Baumeister, 2002). The function of genes in protozoa, and their importance in pathology in relation to the host have been studied using several techniques collectively referred to as reverse genetics (Fig.1.10) (Clayton, 1999). Experiments targeting gene function include transposon mutagenesis that creates mutant genes (Leal *et al.*, 2004; Damasceno *et al.*, 2010), and the transfection of parasites with transient plasmids which is useful for post transcriptional analysis of genes (Clayton, 1999). The ultimate validation of gene function is achieved by either partial or complete disruption of the gene of interest using gene knock-out experiments or by RNA interference

(RNAi) (Clayton, 1999; Baumeister, 2002; Ullu *et al.*, 2002; Militello *et al.*, 2008). Genes of interest can also be overexpressed (gene knock-in) to study their functions in regulatory pathways (Baumeister, 2002; Gardiner *et al.*, 2006).

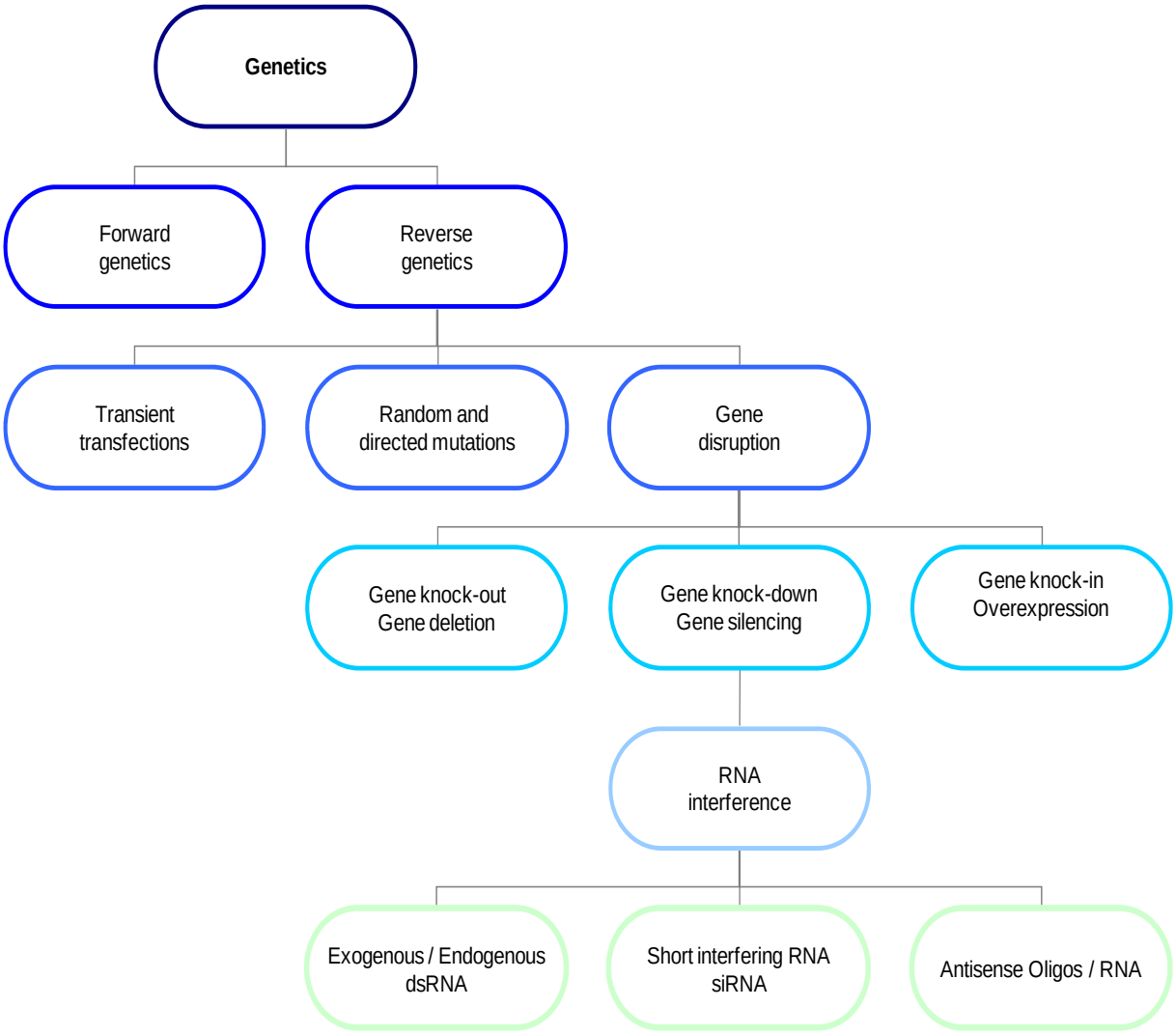


Figure 1.10 Different techniques used to study the function of genes using reverse genetics. Definitive function of a gene is achieved by either partial knock-down of transcribed RNA (RNAi) or complete deletion of DNA (gene knock-out). Overexpression of a gene in a regulatory pathway can also result in a new phenotype.

1.5.1 RNAi in kinetoplastid protozoa

The RNAi mechanism does not exist in all pathogenic parasites (Table 1.3). The first reports of RNAi in protozoa was described in *Trypanosoma spp.* with clear bioinformatic and functional evidence of the pathway in *T. b. brucei* (Ngo *et al.*, 1998; Ullu *et al.*, 2004). These experiments were initially used to characterise structural proteins such as α- and β- tubulin, the silencing of which resulted in the appearance of fat multinucleated cells referred to as

FAT cells (Ngo *et al.*, 1998). This phenotype was attributed to α-tubulin short interfering RNA (siRNA) that disrupted cytokinesis and cytoskeletal organisation of trypanosome cells. Other structural proteins targeted in early experiments were paraflagellar rod protein 2 (PFR2), flagellum attachment protein (FLA) and actin (Ngo *et al.*, 1998; Rusconi *et al.*, 2005). PFR2 was used in studies demonstrating that RNAi in trypanosomes targets specific sequences responsible for the phenotype observed without affecting other genes (Rusconi *et al.*, 2005).

Studies have also explored other targets with less obvious phenotypes than structural proteins. The majority of these are enzymes responsible for various pathways in trypanosomes. In *T. brucei*, enzymes targeted for RNAi include argonaute protein that is involved in processing siRNA during degradation of targeted transcripts (Ngo *et al.*, 1998; Rusconi *et al.*, 2005). Experiments in *T. brucei* targeting acetyl:succinate CoA-transferase (ASCT) was used to confirm the link between the ASCT gene sequence and mitochondrial ASCT activity (Riviere *et al.*, 2004). ASCT is, however, not an essential enzyme as it is not the only acetate producing enzyme in trypanosomes. Other metabolic enzymes that have been characterised using RNAi are iron superoxide dismutase (SOD) (Wilkinson *et al.*, 2006) and trypanothione synthetase (TryS) (Ariyanayagam *et al.*, 2005). Mitochondrial SOD was found to be essential for parasite survival only after RNAi induced parasites were exposed to paraquat, a superoxide generating agent (Wilkinson *et al.*, 2006). A similar trend was observed targeting TryS for RNAi where TryS deficient parasites where sensitive to arsenical drugs indicating a role for trypanothione metabolism in their mode of action (Ariyanayagam *et al.*, 2005).

Several peptidases have also been targeted for RNAi in *T. brucei* parasites. Silencing the major surface peptidase (MSP) in *T. brucei* procyclics decreased the amount of transgenic surface glycoprotein released with a associated increase in cell associated glycoprotein (LaCount *et al.*, 2003). Silencing MSP in bloodstream forms is toxic to the parasite making MSP a valid drug target (LaCount *et al.*, 2003). Cathepsin B-like peptidase, rather than *Tb*CATL (brucipain) was confirmed as the major target in *T. brucei* for the diazomethane inhibitor Z-Phe-Ala-CHN_2 using RNAi (Mackey *et al.*, 2004). RNAi against *Tb*CATL did not result in any obvious phenotype in cultured trypanosomes, whereas silencing the cathepsin B gene led to an enlargement of the endosome, defective cytokinesis and ultimately induced the death of cultured parasites (Mackey *et al.*, 2004). These results were confirmed in a mammalian model where RNAi was induced during infection in mice using doxycycline

(Lecordier *et al.*, 2005). RNAi against cathepsin B in mice infected with parasites bearing cathepsin B gene silencing constructs led to parasite clearance and prevented a lethal infection (Mackey *et al.*, 2004). Silencing *Tb*CATL expression during infection did not cure the mice but reduced parasite virulence with 50% of mice infected with trypanosomes containing *TbCATL* gene silencing constructs surviving for 60 more days than wild type infections (Abdulla *et al.*, 2008). *Tb*CATL was implicated in lower virulence because of its involvement in parasites' ability to cross the blood-brain barrier, thereby confirming previous results using cysteine peptidase inhibitors (Nikolskaia *et al.*, 2006a).

A lysosomal associated membrane protein (LAMP) – like protein 67 (p67) in *T. brucei* was induced for RNAi to assess what role it plays in trafficking *Tb*CATL and human trypanolytic factor (TLF) (Peck *et al.*, 2008). Depletion of RNA transcribing for p67 had no effect on *Tb*CATL trafficking to lysosomes, and endocytosis was unaffected although lysosomes became enlarged after 24 hours of RNAi induction. On the other hand, parasites that were previously sensitive to TLF became resistant to lysis indicating that p67 was essential for TLF activity in lysosomes although a specific role was not established (Peck *et al.*, 2008). Experiments on *T. congolense* showed that the RNAi pathway exists in the parasite. Genes coding for the argonaute and dicer proteins that are involved in the RNAi mechanism have been identified in *T. congolense* (Shi *et al.*, 2006a). Experiments targeting tubulin and luciferase reporter genes for RNAi in *T. congolense* have been successful (Inoue *et al.*, 2002; Coustou *et al.*, 2010). Parasites from the genus stercoraria lack the genes coding for proteins involved in RNAi and experiments in *T. cruzi* do not give expected phenotypes (DaRocha *et al.*, 2004). Other protozoan parasites that lack the RNAi pathway include *Theileria* and *Babesia spp.* and some *Leishmania spp.* (Militello *et al.*, 2008). *Leishmania spp.* in the subgenus *Viannia* have an RNAi pathway. This has been proved with RNAi experiments on *L. braziliensis* targeting green fluorescent protein reporter genes and the argonaute protein (Lye *et al.*, 2010). Proving RNAi in the *Plasmodium spp.* has so far yielded controversial results (Blackman, 2003). Experiments on falcipain-1 and 2, the major cysteine peptidases in *P. falciparum*, using dsRNA demonstrated nuclease activity that led to the fragmentation of RNA (Malhotra *et al.*, 2002). Phenotypes mimicking the addition of the cysteine peptidase inhibitor E-64 were also observed. Recent literature on these studies attributes the results to mechanisms other than RNAi such as general toxicity of exogenous RNA added in the experiment and specific antisense effects that physically obstruct transcription (Baum *et al.*, 2009).

Table 1.3: RNAi experiments conducted in different protozoan parasites [Adapted from Militello *et al.* (2008)]

Organism	Bioinformatic studies	Functional studies	siRNA generation	Endogenous siRNA
T. brucei	+	+	+	+
T. congolense	+	+	+	-
L. braziliensis	+	-	-	-
Giardia intestinalis	+	-	-	-
Entamoeba histolytica	+	-	-	-
Trichomonas vaginalis	+	-	-	-
Toxoplasma gondii	+	-	-	-
Plasmodium spp.	-	-	+	+
T. cruzi	+	-	-	-
L. major	-	-	-	-
L. infantum	-	-	-	-
Cryptosporidium spp.	-	-	-	-
Theileria spp.	-	-	-	-
Babesia bovis	-	-	-	-
Eimeriia tenella	-	-	-	-

RNAi has also been useful when performing high-throughput studies on genes whose functions have not been elucidated, especially with the completion of the *T. brucei* genome sequencing (Berriman *et al.*, 2005). RNAi target sequencing (RIT-seq) uses a plasmid library containing randomly sheared genomic DNA fragments that have been cloned into RNAi vectors (Alsford *et al.*, 2011). Recombinant vectors were then used to transform blood-stream form and insect form parasites and clones that showed a growth defect after three and six days were selected for genomic DNA extraction and sequencing using vector primers. A differentiation defect was also assessed by adding cis-aconitate to transformed bloodstream form parasites and selecting those that had a defect in differentiating to procyclics (Overath *et al.*, 1986). Data acquired from sequencing was mapped using the *T. brucei* genome to identify coding sequences. A total of 1908 and 2724 coding sequences were identified as significant to parasite survival after 3 and 6 days respectively of RNAi induction in the bloodstream form; 1972 and 2677 coding sequences were identified for procyclics and differentiating cells respectively (Alsford *et al.*, 2011). Inducible RNAi libraries have also been

used to identify drug transporters in *T. brucei* bloodstream forms and RNA recognition motifs involved in RNA processing (Wurst *et al.*, 2009; Pillay, 2011).

1.5.1.2 Mechanism of RNA interference

RNA interference (RNAi) has been used as a tool to study gene function and identify possible drug and vaccine targets in trypanosomes (Clayton, 1999). RNAi was first described in *Caernohabditis elegans* when experiments carried out to analyse gene function using antisense approaches produced the same phenotype as the sense control (Guo and Kemphues, 1995). The term 'RNA interference' was described after experiments using double stranded RNA (dsRNA) to silence the *unc2* gene in nematodes were carried out (Fire *et al.*, 1998). It is defined as the mechanism through which gene-specific dsRNA triggers the degradation of homologous transcripts and is a natural phenomenon found in nearly all eukaryotic organisms (Fire *et al.*, 1998; Montgomery *et al.*, 1998; Ullu *et al.*, 2002; Ullu *et al.*, 2004)

There are several natural endogenous roles for RNAi in trypanosomes. This pathway has been implicated in several gene silencing functions such as heterochromatin assembly and maintenance, DNA and histone methylation, DNA elimination, promoter silencing and developmental control (Ullu *et al.*, 2004). *T. brucei* parasites deficient in RNAi have high levels of retrotransposons, indicating that RNAi protects an organism's genome by targeting mobile elements (transposons) (Shi *et al.*, 2004b). RNAi has also been suggested as a natural antiviral defence mechanism since several double stranded viruses that infect protozoa have been identified (Militello *et al.*, 2008). Another possible function for RNAi is its role in regulating gene expression through the action of microRNAs (miRNAs) that inhibit translation by binding to the 3'UTR region of target mRNAs (Sharp, 2001; Ullu *et al.*, 2004; Shi *et al.*, 2004b). miRNAs differ from siRNA in that siRNA are generated from the cleavage of long endogenous or exogenous dsRNA whereas miRNA are naturally occurring regulatory RNAs (Ullu *et al.*, 2004).

RNAi in *T. brucei* relies on the introduction of double-stranded RNA molecules into cells to block the translation of messenger RNA into protein (Fig.1.11). The double-stranded RNA, which must be partly identical in sequence to the gene to be inhibited, is cleaved into shorter fragments called siRNA (25 nt long) by an enzyme called dicer (Shi *et al.*, 2006a) (Fig.1.11). Experiments have shown that there are two forms of dicer in *T. brucei* (*Tb*DCL): cytoplasmic

*Tb*DCL1 that silences potentially harmful viruses in the cytoplasm and nuclear *Tb*DCL2 that is responsible for the *T. brucei* nuclear RNAi pathway (Durand-Dubief and Bastin, 2003). Dicer is a multi-protein comprising of one PAZ domain [named after three proteins containing this domain: Piwi, Argo and Zwille (Cerutti *et al.*, 2000)] that binds RNA and two adjacent RNAse III domains at the N terminal half of the protein (RNAse IIIa and RNAse IIIb) (Fig.1.12). The PAZ domain and the RNAse III domains are 65 angstroms apart and because of this distance, the siRNAs produced in *T. brucei are* 25 nucleotides (nt) long (Shi *et al.*, 2006a).

The siRNA generated by dicer are then split into 'antisense' and 'sense' strands by the argonaute protein, AGO1, which is the catalytic component of the RNA-induced silencing complex. The RNA sense strand is complementary to the DNA template used for transcription and is used as a template for translation and protein synthesis. RNA that is complementary to sense RNA is referred to as antisense RNA. The sense strand is degraded by AGO1 and the remaining antisense strand is incorporated into a protein complex called the RNA-induced silencing complex (RISC) and acts as a guide for the degradation of complementary mRNAs transcribed within the parasite (Shi *et al.*, 2004a; Shi *et al.*, 2006b). The RISC complex is also referred to as slicer (Fig.1.11). Slicer is a member of the argonaute (AGO) family containing an RNAse H motif that cleaves mRNA. *T. brucei* AGO protein (*Tb*AGO1) contains two domains, a PAZ domain similar to the one found in dicer, and a Piwi domain which contains RNAse H-type activity that acts as the active site for slicer (Fig.1.12) (Shi *et al.*, 2004a). In addition to these two domains, *Tb*AGO1 has an N-terminal domain with a high abundance of RGG repeats (Shi *et al.*, 2004a). The PAZ domain functions as an anchoring site for the two-nucleotide 3'-overhang of siRNAs (Shi *et al.*, 2004a). The activity of slicer is directed by the antisense RNA towards the target mRNA, which is degraded, thus effectively 'silencing' the gene from which it was transcribed (Novina and Sharp, 2004).

Initial experiments using dsRNA in *T. brucei* involved the transient transfer of synthetically synthesised dsRNA (Ngo *et al.*, 1998). RNA interference induced in this manner was temporary and not transferable to the next generation. Methods were adapted from work in *C. elegans* that involved soaking worms in dsRNA solution instead of injecting them to induce RNAi (Tabara *et al.*, 1998). This method proved to be inefficient due to its transient nature and alternatives were developed. Heritable and inducible vectors for RNAi provide a more stable option because they are integrated into the parasite genome and can be transferred from one generation to the next. They also contain a tetracycline repressor which enables

RNAi to be induced by the addition of tetracycline to the growth medium of transfected parasites (Sui *et al.*, 2002).

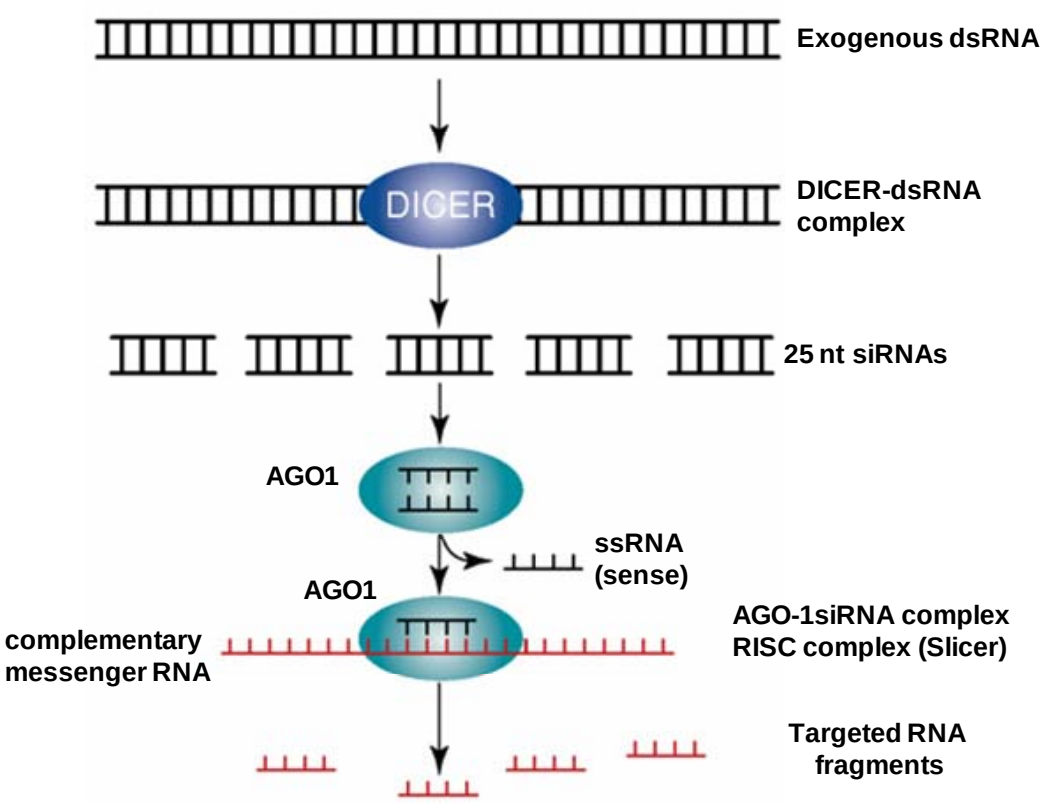


Figure 1.11 Schematic representation of the RNAi system in trypanosomes. Dicer cleaves exogenous dsRNA into 25 nt long siRNA which is separated by AGO1 protein into sense and antisense strands. The antisense strand then guides the AGO1-siRNA complex to degrade complementary mRNAs causing silencing of RNA from which the original dsRNA was transcribed (Balaña-Fouce and Reguera, 2007).

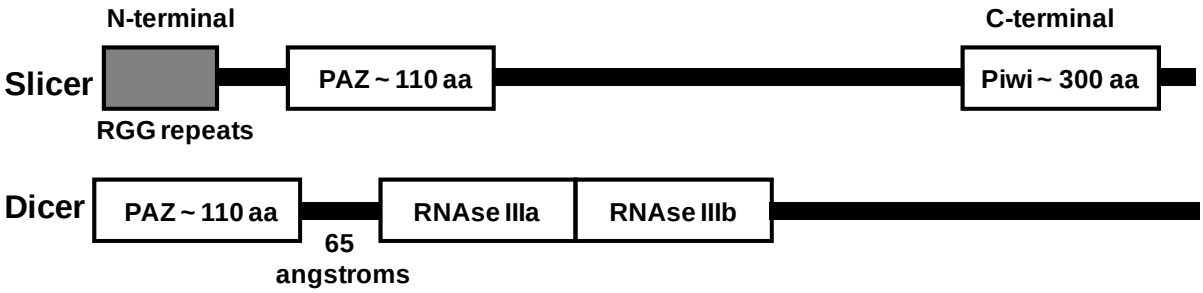


Figure 1.12 Structures of dicer and slicer proteins. Both proteins possess a PAZ RNA binding domain of approximately 110 amino acids in size that is responsible for anchoring dsRNA for processing. Slicer also possesses a Piwi domain with RNAse H activity for degrading complementary mRNA. The distance between the PAZ domain and RNAse III domains in dicer is 65 angstroms accounting for the 25 nt long siRNA generated.

Two vectors for RNAi were constructed by modifying the backbone vector pLew 100 (Fig.1.12) (Wirtz *et al.*, 1999). Both of these vectors use tetracycline inducible promoters to ensure the stable expression of dsRNA in transfected parasites. The hairpin vector utilises the *pLew 79* vector backbone (Wirtz *et al.*, 1999) in which the luciferase gene is substituted by a cassette consisting of two inverted repeats of the mRNA target sequence separated by a stuffer sequence. RNA transcription is driven by the tetracycline inducible PARP promoter and produces a hairpin like RNA molecule, also referred to as stem-loop RNA (Shi *et al.*, 2000).The double promoter vector on the other hand utilises the *pLew 82* vector backbone (Wirtz *et al.*, 1999) and has two opposing tetracycline-inducible T7 RNA promoters that flank the mRNA target sequence (Shi *et al.*, 2000). Examples of RNAi vectors that utilise dual T7 RNA polymerase promoters are pZJM, that contains a hairpin vector cassette (Fig.1.13) (Wang *et al.*, 2000) and the p2T7^{Ti} vector that contains a double promoter vector cassette (Fig.1.13) (LaCount *et al.*, 2002). Double promoter vectors have also been used in other trypanosome species for RNAi. The p2T7^{Ti} vector has been adapted in *T. congolense* with modifications that target integration into the β-tubulin intergenic region.

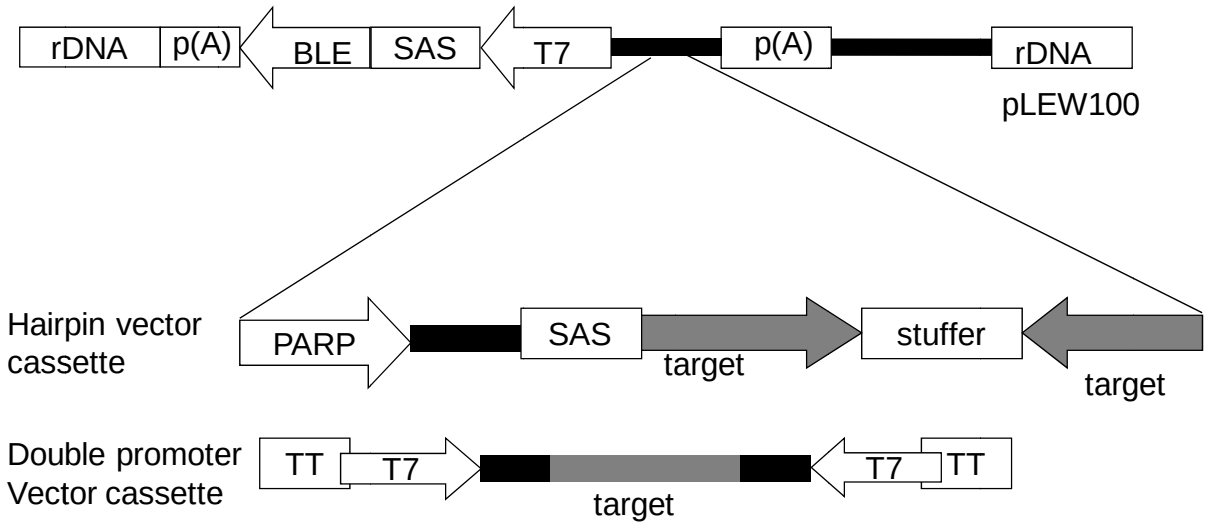


Figure 1.13 Hairpin and double promoter vectors for achieving stable RNAi responses in *T. brucei*. Tetracycline-inducible PARP promoter (PARP); tetracycline operator regions ( ); T7 transcription terminators (TT); T7 promoters (T7); ribosomal DNA non-transcribed spacers (rDNA); poly (A) addition sites (p(A)); phleomycin resistance genes (BLE) and splice-site acceptor (SAS) [Adapted from Wirtz *et al.* (1999)].

Additional plasmids containing a T7 promoter (pLEW 29) and a tetracycline repressor gene (pLEW 13) are also used to transform parasites in conjunction with the p2T7^{Ti} vector for

functional RNAi silencing in *T. congolense* (Inoue *et al.*, 2002). Integration of a stem and loop vector into *T. cruzi* was successful but failed to induce RNAi due to the absence of an RNAi mechanism in the parasites (DaRocha *et al.*, 2004).

1.5.2 Gene knock-out studies in kinetoplastid protozoa

Gene knock-out techniques involve the deletion of DNA coding for the gene under study. This requires the transfection of parasites with replacement plasmid vectors that target the gene of interest for deletion (Clayton, 1999). Because kinetoplastid parasites are diploid, two rounds of transfection are required to delete both copies of the targeted gene. Drug markers that confer resistance are used to select for recombinant heterozygotes (with one copy of the gene deleted) and null mutants (with both copies of the gene deleted). Antibiotic resistance markers routinely used when transforming trypanosomes and *Leishmania* include puromycin acetyltransferase, neomycin phosphotransferase, hygromycin phosphotransferase and bleomycin resistance protein (Table 1.4) (Clayton, 1999; Ommen *et al.*, 2009).

In kinetoplastid parasites where the RNAi mechanism is absent or ineffective (*Leishmania spp.* and *T. cruzi*), gene deletion studies have taken precedence over other reverse genetics techniques (Xu *et al.*, 2009). The first study on gene deletion in *Leishmania* was the knock-out of the gene coding for dihydrofolate reductase-thymidylate synthase (DHFR-TS) in *L. major* (Cruz *et al.*, 1991). The replacement strategy used was sequential with the replacement of one gene to create a heterozygote, followed by the generation of null mutants (Fig.1.13). DHFR-TS null mutants grew only in the presence of thymidine, suggesting that the sole role of DHFR-TS was the provision of thymidine in *L. major* (Cruz *et al.*, 1991). Early experiments on transformation of DNA in *T. brucei* evaluated the viability of inserting DNA plasmids into intergenic regions of the αβ-tubulin gene and ribosomal RNA (rRNA) (Lee and Van der Ploeg, 1990; Carruthers *et al.*, 1993). The highest transformation efficiencies were observed when targeting rRNA and using Zimmerman post-fusion medium with glucose (ZPFMG) (Carruthers *et al.*, 1993). Knocking-out the thioredoxin gene in *T. b. brucei* bloodstream and procyclic parasites had no effect on their viability (Schmidt *et al.*, 2002). A possible explanation for the lack of a phenotype was the abundance of tryparedoxin, an enzyme that has lower but overlapping specificities with thioredoxin in *T. b. brucei*.

Table 1.4: Markers available for selecting permanently transformed trypanosomes and *Leishmania*.

Selectable marker	Selective Drug	Drug target	Mode of resistance
Neomycin phosphotransferase (NEO)	G418	Ribosome	Enzymatic inactivation of drug[a,b]
Blasticidin S deaminase (BLA)	Blasticidin S	Ribosome	Enzymatic inactivation of drug[c,d]
Hygromycin phosphotransferase (HYG)	Hygromycin	Ribosome translocation	Enzymatic inactivation of drug[e,f]
Bleomycin resistance protein (BLE)	Phleomycin, zeocin	DNA (causes breakage)	Binding (sequestration) of drug[g,h]
Puromycin acetyltransferase (PAC)	Puromycin	Protein synthesis	Enzymatic inactivation of drug[h,i]
Streptothricin acetyltransferase (SAT1)	Nouseothricin	Ribosome	Enzymatic inactivation of drug[j,k]

[a]Laban *et al.* (1990), [b]ten Asbroek *et al.* (1990), [c]Kobayashi *et al.* (1991), [d]McCulloch and Barry (1999), [e]Cruz *et al.* (1993), [f]Lee and van der Ploeg (1991), [g]Jefferies *et al.* (1993), [h]Freedman and Beverley (1993), [i]Lorenz *et al.* (1998), [j]Ruepp *et al.* (1997), [k]Joshi *et al.* (1995).

Knocking-out the same gene in two different, but related parasites can lead to different phenotypes observed. GPI: protein transamidase complex null mutants (∆GP18) in *T. brucei* were unable to establish infections in the tsetse midgut (Lillico *et al.*, 2003). Knocking out the same gene in *L. mexicana* did not change the virulence and survival of ∆GP18 null mutant parasites in the mammalian host (Hilley *et al.*, 2000). It is also possible to knock-out several gene repeats in one reaction. Leishmanolysin, a surface metallopeptidase of 63 kDa (gp63), is found in a gene array with seven repeats that spans a 20 kb region (Joshi *et al.*, 2002). ∆gp63 mutants were generated by deleting this entire region and results showed that gp63 was a virulence factor in *L. major*. Cysteine peptidases in *L. mexicana* are encoded by a tandem array of 19 genes in the *lmcpb* locus and two single copy genes, *lmcpa* that is cathepsin L-like, and *lmcpc* that is cathepsin B-like (Mottram *et al.*, 1996). The disruption of the entire *lmcpb* gene array along with the *lmcpa* single copy cathepsin L-like genes was successfully achieved using gene knock-out techniques with null mutant parasites showing reduced virulence in Balb/c mice (Mottram *et al.*, 1996). The reinsertion of amastigote-specific and metacyclic-specific cysteine peptidase genes into *lmcpb* null mutant parasites did not

improve virulence which was only restored with the re-expression of the multiple *lmcpb* genes (Mottram *et al.*, 2004).

It is not possible to obtain null mutant clones when targeting genes that are essential to the parasite under study. Amastigote specific cysteine peptidase in *L. chagasi* is encoded by a single copy gene (Mundodi *et al.*, 2002). Knock-out studies were only able to achieve heterozygotes with the reduced ability to infect macrophage cells (Mundodi *et al.*, 2005). Gene deletion does not always result in an observable phenotype (Estevez *et al.*, 1999; Schmidt *et al.*, 2002). Deleting the gene coding for glutamate dehydrogenase (GDH) did not affect the abundance of edited RNAs, a function that has been attributed to GDH (Estevez *et al.*, 1999). It is also important to observe how other genes in the parasite are affected by the deletion of one gene (Deutscher *et al.*, 2008; Swenerton *et al.*, 2011). Deletion of oligopeptidase B in *L. donovani* led to a large up-regulation of enzymatically inactive membrane associated enolase. It could be inferred that OPB clears membrane associated enolase in the parasite to help evade macrophage activation during early stages of infection (Swenerton *et al.*, 2011).

The common method used to delete the two alleles of a gene in trypanomastids is to flank an antibiotic resistance marker gene with the 5'- and 3'- intergenic regions of the target gene (Fig.1.13) (Ommen *et al.*, 2009). For specific and efficient gene targeting, up to 1000 base pairs of both flanking regions can be cloned into the knock-out vector (Fig.1.14). Recombinant vectors are then transfected into parasites where homologous recombination is responsible for replacing the targeted gene with a selection marker. One step gene replacement has been suggested in *Leishmania spp.* (Ommen *et al.*, 2009). In this experiment, both genes are replaced in one reaction making it a viable alternative to first creating heterozygotes. A polymerase chain reaction-based strategy has been used in *T. brucei* (Gaud *et al.*, 1997). This involves generating a knock-out vector using the 5' and 3' DNA regions flanking the gene of interest and drug resistance genes as templates in one PCR reaction thereby drastically reducing the amount of work required for cloning (Gaud *et al.*, 1997; Xu *et al.*, 2009). The multi-site gateway system is even more efficient than the PCR based system and involves using commercial vectors that have homologous recombination sites that do not require any cloning reactions to incorporate flanking regions (Xu *et al.*, 2009).

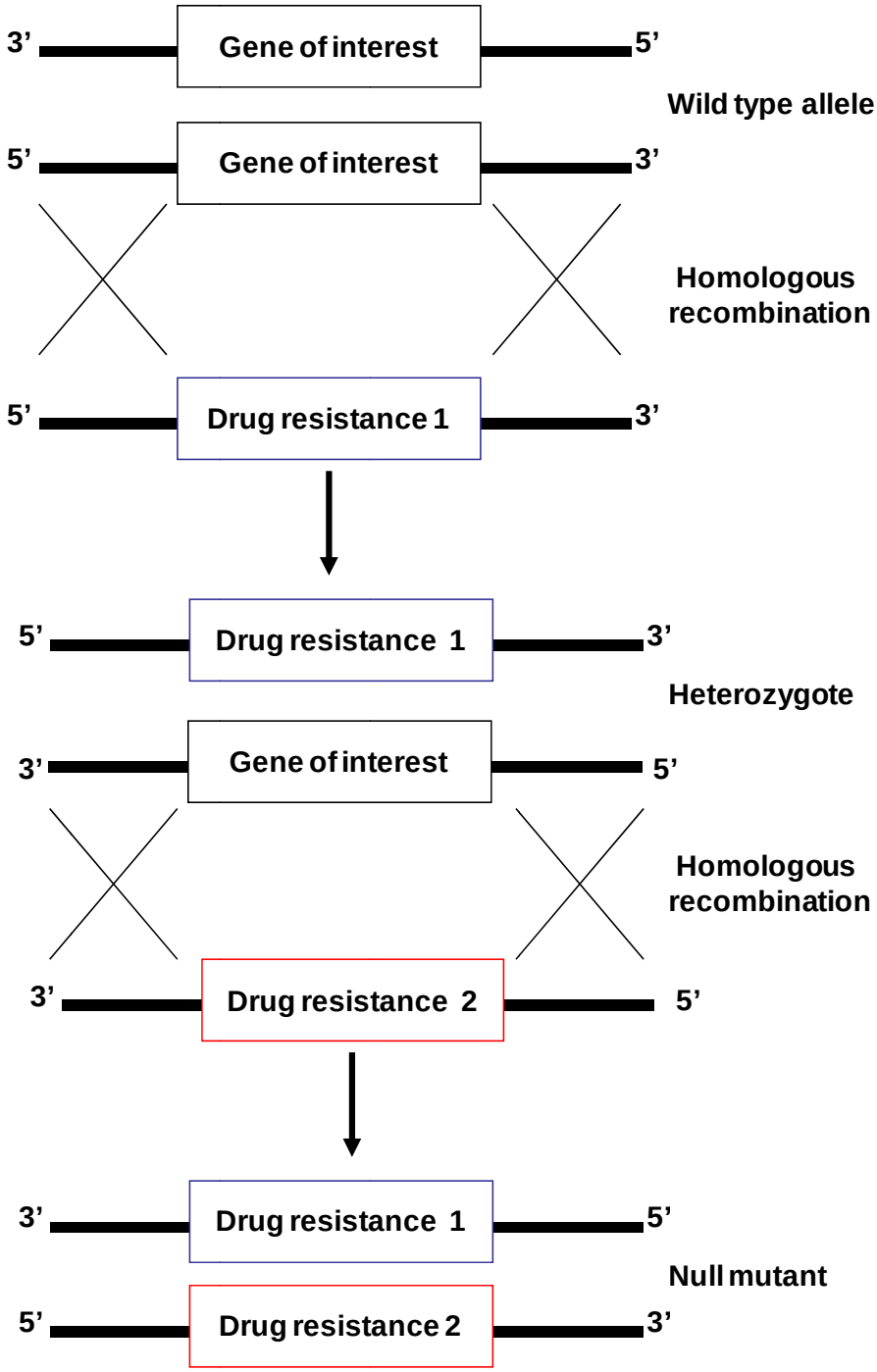


Figure 1.14 Generation of null mutants by homologous recombination. Protozoan parasites are diploid and therefore need two gene replacement reactions using two different drug resistance markers for each copy of the gene to be deleted. The order of replacement does not have an effect on the process. Parasites with one gene replaced are referred to as heterozygotes, and parasites with both genes replaced are referred to as null mutants. 5' ▬ 3' represent the flanking regions of the gene to be deleted and are used as targets for homologous recombination.

1.5.2.1 Mechanism of homologous recombination

Homologous recombination is the process in which nucleotide sequences are exchanged between two identical sequences of DNA and is used to repair double stranded DNA breaks and ensuring the completion of DNA replication (Kanti Bhattacharyya *et al.*, 2004; Sung and Klein, 2006). Although double stranded DNA breaks are considered to be dangerous to the cell, they can be introduced intentionally, the best example being during meiosis to facilitate genetic exchange (Proudfoot and McCulloch, 2006). In trypanosomes, homologous recombination is one of the mechanisms involved in antigenic variation during telomere exchange that switches a silent VSG site with an active VSG expression site (Taylor and Rudenko, 2006). Homologous recombination was first described in *E. coli* where the RecA protein was found to be central to the process (Cox, 2003). Rad51 and Dmc1 (disrupted meiotic cDNA) proteins are the eukaryotic homologs of RecA and are responsible for binding to DNA, searching for homologous DNA molecules and subsequently exchanging counterpart DNA strands (Fig. 1.15) (Kanti Bhattacharyya *et al.*, 2004). Rad51 is responsible for recombination events during mitosis between sister chromatids and requires the involvement of Dmc1 for recombination between two different parental chromosomes during meiosis (Masson and West, 2001). Dmc1 is therefore meiosis specific and although single copy genes have been identified in *T. brucei*, *T. cruzi* and *L. major*, Dmc1 mutation does not result in any alterations to DNA repair and recombination (Proudfoot and McCulloch, 2006).

Rad51 has been identified in both *T. brucei* and *L. major* (McCulloch and Barry, 1999; McKean *et al.*, 2001). Disruption of the *Rad51* gene in *T. brucei* down-regulates DNA recombination and VSG switching (McCulloch and Barry, 1999), and exposure of *L. major* parasites to phleomycin (Table 1.4) up-regulates its expression (McKean *et al.*, 2001). Five *Rad51*-related genes have been identified in *T. brucei* and mutations in two of them, *Rad51-3* and *Rad51-5*, resulted in reduced DNA recombination and rendered parasites sensitive to DNA damage by methyl methanesulfonate (Proudfoot and McCulloch, 2005). Rad51-3 was also found to contribute to VSG- switching (Proudfoot and McCulloch, 2005). Another gene product that has been found to be associated with Rad51 in *T. brucei* is *BRAC2* (breast cancer susceptibility gene) (Hartley and McCulloch, 2008). BRAC2 regulates homologous recombination by binding to Rad51 and blocking its activity until it is needed for repair (Hartley and McCulloch, 2008). Although the Rad51 pathway (Fig. 1.15) is dominant during the repair of double stranded breaks in trypanosomes, other mechanisms that have been implicated in DNA repair are the involvement of the Mre11 complex (Robinson *et al.*, 2002;

Tan *et al.*, 2002) and mismatch repair (Barnes and McCulloch, 2007). Rad51 mediated homologous recombination begins with the binding of single stranded DNA to single strand binding protein that is followed by replacement with Rad52 (Fig. 1.15) (Kanti Bhattacharyya *et al.*, 2004). Rad52 directs Rad51 to bind to the single stranded DNA forming a nucleoprotein filament. Each molecule of Rad51 binds to three nucleotides of single stranded DNA forming a right handed helix. Rad54 then associates with Rad51 in the helix and is responsible for ATPase activity that drives DNA strand exchange (Kanti Bhattacharyya *et al.*, 2004).

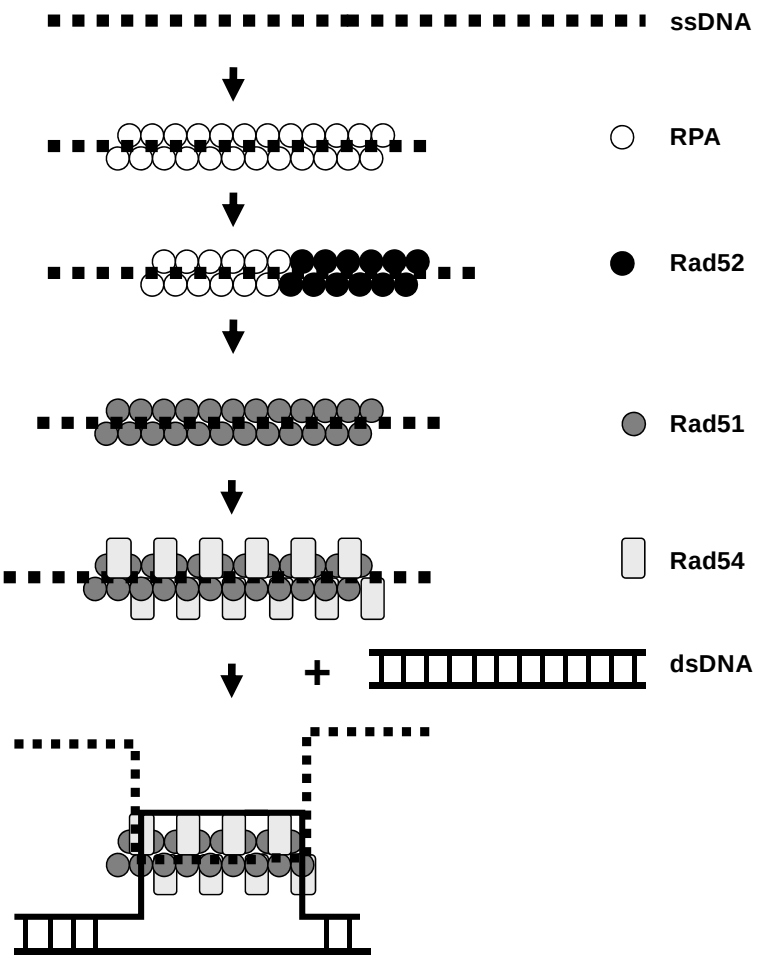


Figure 1.15 The Rad51 pathway for homologous recombination. Single strand binding protein (RPA) binds single stranded DNA (ssDNA) and is replaced by Rad52 that directs the attachment of Rad51 to ssDNA. Rad54 associates with Rad51-ssDNA helix and provides ATPase activity that is required for strand exchange with double stranded DNA (dsDNA) [Adapted from Kanti Bhattacharyya *et al.* (2004)].

1.6 OBJECTIVES AND OUTLINE OF THE PRESENT STUDY

The main aim of this study was to investigate the function of two trypanosome peptidases that have been identified as pathogenic factors during infection, namely *Tco*CATL and oligopeptidase B (Authié, 1994; Coetzer *et al.*, 2008).

The objectives identified to achieve the main aim of the study were:

- The induction of RNAi targeting *Tco*CATL in *T. congolense* parasites using the p2T7^{Ti} vector in order to assess the *in vitro* functions of the peptidase
- To generate *Tco*CATL gene array knockout *T. congolense* parasites that would be used for mouse infections so as to study what role *Tco*CATL plays during infection and its function as a pathogenic factor
- The induction of RNAi targeting OPB in *T. congolense* and *T. brucei* parasites using the p2T7^{Ti} vector for analytical *in vitro* studies that would assess the importance of the peptidase for parasite division
- The generation of OPB null mutants in *T. congolense* and *T. brucei* for *in vivo* studies in mice that would confirm previous *in vitro* studies used to define the role of OPB during infection as a pathogenic factor.

Functional analysis of these two enzymes was achieved using RNAi that down-regulates mRNA transcribing for the genes of interest and gene knock-out that deletes the genes at DNA level. Oligopeptidase B was studied in the *T. congolense* procyclic forms using RNAi and in *T. brucei* bloodstream and procyclic forms using both RNAi and gene knock-out. *Tco*CATL was studied in the *T. congolense* bloodstream form in mice following a model that has been performed using *Tb*CATL, the major cysteine peptidase in *T. brucei* (Abdulla *et al.*, 2008).

The *in vivo* down regulation of *Tco*CATL using RNAi is described in Chapter 2. This required the identification of regions in the gene that are conserved across different trypanosome species and amplification of selected regions from genomic DNA. *Tb*CATL from *T. brucei* was also included in the experiment as a positive control. PCR products obtained were cloned into the p2T7^{TI} RNAi vector and recombinant plasmids used to transform a procyclic *T. congolense* inducible strain. A recombinant RNAi vector containing tubulin was used as a positive control for transformation (Inoue *et al.*, 2002). Because *Tco*CATL expression is very low during the procyclic stage, selected clones targeted for *Tco*CATL RNAi were

differentiated into bloodstream forms and used to infect mice where RNAi was induced. Down-regulation was assessed using activity assays specific for *Tco*CATL. The generation of *Tco*CATL knock-outs in *T. congolense* was also attempted and regions flanking the *Tco*CATL gene array were successfully identified using Southern blot and sequencing.

The *in vitro* down regulation of oligopeptidase B in *T. congolense* and *T. brucei* using RNA interference and generation of chicken anti-oligopeptidase B antibodies used to analyse the down-regulation of OPB is described in Chapter 3. Conserved regions in the gene were identified, cloned and transformed into respective parasite strains as described for *Tco*CATL in Chapter 2. Because oligopeptidase B is expressed at adequate levels in the procyclic form, it was not necessary to differentiate transformed parasites into bloodstream forms. RNAi was monitored using reverse transcription and quantitative PCR.

The generation of oligopeptidase knock-outs in *T. brucei* bloodstream parasites is described in Chapter 4. Part of this work was carried out in the laboratory of Prof. Theo Baltz, Laboratoire de Microbiologie Cellulaire et Moléculaire et Pathogénicité, University Victor Segalen Bordeaux2, France as part of the South Africa (NRF) -France Scientific Cooperation research grant. Oligopeptidase B null mutants were confirmed using Southern and western blotting and in activity assays. Additional *in vitro* characterisation of null mutants was performed using microscopy and zymograms. *In vivo* characterisation of null mutants during mouse infections was also studied using activity assays and immunohistochemistry. Part of this work was accepted for publication (Kangethe *et al.*, 2011).

The results and conclusions pertaining to this study are discussed in Chapter 5.

*Tco*CATL, *Tb*CATL and OPB alignments are displayed in the appendix along with a publication generated from this study.

CHAPTER 2

IN VIVO DOWN-REGULATION USING RNA INTERFERENCE AND THE ATTEMPTED GENERATION OF KNOCK-OUTS OF *TCO*CATL, THE MAJOR CYSTEINE PEPTIDASE IN *TRYPANOSOMA CONGOLENSE*

African bovine trypanosomosis, also referred to as nagana, is a parasitic disease that puts large numbers of cattle at risk in sub-Saharan Africa with estimated losses of up to five billion dollars annually (Swallow, 2000; Antoine-Moussiaux *et al.*, 2009). Various methods have been employed to control the disease in cattle with chemotherapy the most commonly used mode of treatment (McDermott and Coleman, 2001). However, there has been a large increase in trypanosome resistance to most drugs (Delespaux and de Koning, 2007), and this has led to the exploration of other avenues for disease control, such as vaccine development (Antoine-Moussiaux *et al.*, 2009). The first vaccines targeted the variable surface coat of the trypanosome parasite, but these only provide protection against homologous infections (Taylor and Mertens, 1999). There is therefore a need to target invariant antigens across heterologous infections when developing an effective vaccine (Authié *et al.*, 1993b; Taylor, 1998). Parasites lysed by the host immune system release a multitude of trypanosome products many of which are peptidases previously confined to the cytoplasm and lysosomes of live trypanosomes (Tizard *et al.*, 1978). Trypanosome peptidases released during infection present an attractive target for the development of vaccines and chemotherapeutic agents as they are invariant (Taylor and Authié, 2004), and are also responsible for a variety of pathologies associated with the disease (McKerrow *et al.*, 2006).

Papain-family C1 cysteine peptidases, belonging to clan CA of cysteine peptidases, are released by trypanosomes during infection (Taylor and Authié, 2004; McKerrow *et al.*, 2006; Antoine-Moussiaux *et al.*, 2009). Family C1 is further divided into two major groups, cathepsin L (CATL)-like and cathepsin B (CATB)-like which in *T. brucei* are referred to as *Tb*CATL and *Tb*CATB respectively according to recently suggested nomenclature (Caffrey and Steverding, 2009). *CATL* genes in kinetoplastid parasites occur in multiple copies with *T. brucei* CATL having more than 20 copies arranged in a long tandem array on the same chromosome (Mottram *et al.*, 1989). *T. cruzi* CATL is encoded by more than 130 genes that are arranged in clusters found on two to four different chromosomes (Campetella *et al.*, 1992; Caffrey and Steverding, 2009). *L. chagasi* and *L. donovani* both have five genes encoding CATL; also arranged in a tandem array with other *Leishmania* CATL (*LmCATL*) containing tandem arrays

of 8-20 genes (Caffrey and Steverding, 2009). In *T. congolense*, *CATL* genes (*TcoCATL)* exist as two kb gene units organised in a tandem array consisting of a 1332 bp pre-pro-enzyme coding region, and a highly conserved intergenic region of approximately 660 bp (Kakundi, 2008). The sequence identity between these different copies of *TcoCATL* genes varies from 86% to 99% (Kakundi, 2008; Pillay *et al.*, 2010). Multiple isoforms of *CATL* genes exist in *L. mexicana* (Mottram *et al.*, 1997), *T. cruzi* (Lima *et al.,* 2001) and *T. congolense* (Fish *et al.*, 1995; Boulangé *et al.*, 2001). *Lm*CATL isoenzymes are expressed at different life cycle stages of the parasite, and have different substrate specificities (Mottram *et al.*, 1997). *Tco*CATL gene expression is up-regulated at the bloodstream form stage of the parasite and although the *Tco*CATL isoenzymes CP1 and CP2 (congopain), have similar sequences (90% identical), they possess fundamental functional differences suggesting distinct roles for the enzymes *in vivo* (Boulangé *et al.*, 2001).

CATL and CATB have been implicated as virulence factors involved in parasite immunoevasion as well as cell and tissue invasion (Sajid and McKerrow, 2002) and have been evaluated as targets for various chemotherapeutic agents and inhibitors (Troeberg *et al.*, 1999; Troeberg *et al.*, 2000; Lecaille *et al.*, 2002). In *T. brucei* infections in mice, CATB has been implicated in the degradation of host proteins and pathology (Mackey *et al.*, 2004; Abdulla *et al.*, 2008). Targeting *Tb*CATB rather than *Tb*CATL *in vitro* for RNAi led to endosome enlargement, defective cytokinesis after mitosis and death of parasites (Mackey *et al.*, 2004). *In vivo* experiments targeting *Tb*CATB and *Tb*CATL using RNAi showed that the absence of *Tb*CATB led to the clearance of parasites from the bloodstream when compared to controls which died within 13 days (Abdulla *et al.*, 2008). Fifty per cent of the mice used in the experiment that targeted *Tb*CATL for RNAi survived 60 days longer compared to controls (Abdulla *et al.*, 2008). *Tb*CATL was implicated in assisting parasite entry into the brain when using an *in vitro* model of the human blood-brain barrier (Abdulla *et al.*, 2008).

*Tco*CATL is the major cysteine peptidase in *T. congolense* bloodstream form parasites (Authié *et al.*, 1992) and studies relating to its functions during infection have mainly concentrated on the role it plays in immunity (Authié *et al.*, 1992; Authié *et al.*, 1993a; Authié *et al.*, 1993b; Authié, 1994; Authié *et al.*, 2001). *Tb*CATL has been implicated in assisting parasites cross the human blood-brain barrier (Abdulla *et al.*, 2008), but assigning a similar function to *Tco*CATL in *T. congolense* would be difficult given that *T. congolense* parasites are exclusively intravascular and do not leave the bloodstream of the host (Ojok *et al.*, 2002).

To study what role *Tco*CATL plays in bloodstream form *T. congolense* physiology and host pathogenesis, RNAi knock-down parasites were generated and used in a mouse model to tract infection of parasites deficient in *Tco*CATL.

For successful RNAi of *CATL* gene products in trypanosomes, it was necessary to induce RNAi during the bloodstream form stage of the parasite as opposed to the procyclic stage. This is because CATL activity is highest at the bloodstream stage of the lifecycle in both *T. brucei* and *T. congolense* parasites (Mbawa *et al.*, 1991). For *T. congolense*, procyclic parasites were first transformed with the recombinant RNAi vector before converting the procyclic cultures into bloodstream parasites for infection of mice. This is because it is difficult to carry out the *in vitro* culture of *T. congolense* parasites and this technique has only recently been perfected (Coustou *et al.*, 2010). *Tco*CATL RNAi transformed *T. congolense* bloodstream parasites were then used to infect mice and *in vivo* RNAi was induced using doxycycline as has been previously performed (Abdulla *et al.*, 2008). An attempt was also made to knock-out the entire gene array of *Tco*CATL and *Tco*CATL-like genes. This required the identification of the regions flanking the *Tco*CATL tandem genes by Southern blot and sequencing and subsequently cloning these regions into knock-out vectors for transformation into *T. congolense*.

2.1 MATERIALS AND METHODS

2.1.1 Materials

Molecular biology: ApaI, NotI, XbaI [reference for nomenclature see Roberts *et al.* (2003)], shrimp alkaline phosphatase (SAP), T4 DNA ligase and ligation buffer, 10 mM dNTP mix, X-gal, IPTG, MassRuler™ DNA ladder mix, GeneJET™ plasmid miniprep kit, TransformAid™ bacterial transformation kit and Biotin DecaLabel™ DNA labeling kit were obtained from Fermentas (Vilnius, Lithuania). pGEM-T® vector was obtained from Promega (Madison, WI, USA). DNA clean and concentrator kit was obtained from ZymoResearch (Orange, CA, USA), E.Z.N.A® gel extraction kit from PEQlab (Erlangen, Germany) and *Escherichia coli* JM 109 cells purchased from New England Biolabs (Ipawitch, MA, USA). FIREpol® Taq polymerase, 10 x PCR reaction buffer and 25 mM $MgCl_2$ were acquired from Solis Biodyne (Tartu, Estonia). Hybond™-N DNA binding membrane was obtained from Amersham (Buckinghamshire, UK). Molecular biology grade agarose was purchased from Conda laboratories (Madrid, Spain), ampicillin sodium salt from USB Corporation (Cleveland, OH,

USA), bacteriological agar and tryptone from Merck (Darmstadt, Germany). The yeast extract was purchased from Sigma (St. Louis, MO, USA). The recombinant p2T7^{Ti}/BIP RNAi plasmid was a kind gift from Professor Theo Baltz (University of Victor-Segalen, Bordeaux 2, France). Recombinant cosmids containing *T. congolense* genomic DNA (strain IL3000) were provided by A. Boulangé (University of KwaZulu-Natal).

Trypanosome culture: Minimum essential medium (MEM), Iscove's Modified Dulbecco's Medium (IMDM), HEPES, sodium pyruvate, thymidine, adenosine, hypoxanthine, adenosine, bathocuprone sulfate acid, glutamine, 2-mercaptoethanol, hemin, proline, cis-aconitate, tetracycline, doxycycline hyclate and cyclophosphamide were obtained from Sigma (St. Louis, MO, USA). Heat inactivated goat serum and heat inactivated foetal calf serum (FCS) were obtained from Gibco (Paisley, UK). Filters (0.2 µm) were purchased from Pall (Ann Arbor, USA). 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 H solution 2 were obtained from Lonza (Levallois-Perret, France). *T. congolense* TRUM183:29-13 and *T. brucei* T7T-29 RNAi inducible parasite strains, and NOD/SCID mice were a gift from Professor Theo Baltz (University of Victor-Segalen, Bordeaux 2, France). Balb/C mice were obtained from the Bioresources Unit (BRU) Westville campus, University of KwaZulu-Natal (UKZN). Protocols using BALB/C mice were approved by the UKZN animal ethics committee (References 036/09/Animal and 069/10/Animal). Protocols using NOD/SCID mice were approved by the University of Bordeaux 2 animal care and use committee and the commission de genie genetique (Direction Generale de la Recherche et de l'Innovation).

Peptide substrate: Benzyloxycarbonyl (Z)-Phe-Arg-7-amino-4-methylcoumarin (AMC) was obtained from Bachem (Torrance, CA, USA). Peptidase activity was measured using a FLUOstar optima spectrofluorometer (BMG Labtech, Offenburg, Germany).

2.2.2 Methods

2.2.2.1 Cloning conserved regions of *Tco*CATL and *Tb*CATL genes for RNAi

A pellet from 1 x 10^7 *T. congolense* (strain IL3000) or control *T. b. brucei* (strain Lister 427) parasites (Mackey *et al.*, 2004; Abdulla *et al.*, 2008), was used to extract genomic DNA as previously described (Medina-Acosta and Cross, 1993). Briefly, the pellet of parasites was resuspended in 1 ml of PBS (pH 7.2), centrifuged (1000 *g*, 10 min, RT), and the supernatant

removed. The pellet was dissolved in 150 µl TELT [50 mM Tris-HCl buffer, pH 8, 62.5 mM EDTA, pH9, 2.5 M LiCl, 4% (v/v) Triton X-100] and incubated for 5 min at RT. Phenol-chloroform [150 µl; 1:1 (v/v)] was added and agitated for 5 min at RT. The aqueous and non-aqueous phases were separated by centrifugation (13 000 *g*, 5 min, RT), and the upper phase added to 300 µl of absolute ethanol and centrifuged (13 000 *g*, 5 min RT) to precipitate genomic DNA. The supernatant was poured out and the DNA pellet washed with 1 ml of absolute ethanol before a last round of centrifugation (13 000 *g*, 5 min, RT). The ethanol was poured out and the pellet of genomic DNA incubated at 37ºC for 15 min to evaporate excess ethanol. The dry DNA pellet was dissolved in 100 µl of TE buffer [100 mM Tris-HCl buffer, pH 7.5, 10 mM EDTA] containing 2 mg/ml RNAse and incubated at 37ºC for 45 min.

DNA sequences coding for the CATL peptidase regions that are conserved between *T. congolense* [EMBL accession no. L25130, (Jaye *et al.*, 1994)] and *T. brucei* [EMBL accession no. AJ297265, (Caffrey *et al.*, 2001)] (see appendix 1) were selected to design primers that amplified a 320 bp fragment in *T. congolense* and a 300 bp fragment in *T. brucei.* The primers were designed to introduce an XbaI restriction site at each end of the PCR product for subcloning into the p2T7^{Ti} RNAi plasmid. Primers for the 320 bp region in the *Tco*CATL region were: forward, (5'-GTGTCTAGA**ATG**GCCGTTGCGG CGTGCTTT-3') with an XbaI restriction site (underlined) and a start codon (bold); reverse, (5'-CTCTCTAGA**TTA**CCCAGT GGACACATTCAC-3') with an XbaI restriction site (underlined) and a stop codon (bold). Primers for the 300 bp region in the *Tb*CATL region were: forward, (5'-GTGTCTAGA**ATG**GCTATGGCAGCGTGCCTT-3') with an XbaI restriction site (underlined) and a start codon (bold); reverse, (5'-CTCTCTAGA**TTA**GCCAGTGGTTACGTTCAC-3') with an XbaI restriction site (underlined) and a stop codon (bold). Extracted genomic DNA was used as a template for PCR. Briefly, the master mix for the PCR reaction contained sets of primers designed for *Tb*CATL or *Tco*CATL (0.25 µM each) with *T. brucei* and *T. congolense* DNA respectively, 1 x PCR buffer, 2.5 mM $MgCl_2$, 1 U Taq and 0.5 mM dNTPs in a total reaction volume of 25 µl. PCR amplification was started with a Taq polymerase activation step at 95ºC for 5 min, followed by 25 cycles of a DNA denaturing step at 95ºC for 30 s, a primer annealing step at 55ºC for 30 s, and an elongation step at 72ºC for 1 min. A final elongation step was added at the end of 25 cycles at 72ºC for 7 min. Each of the PCR products generated (5 µl) were analysed on a 1% (w/v) agarose gel in 1 x Tris-acetate-EDTA (TAE) buffer [40 mM Tris-HCl buffer pH 7.4, 20 mM glacial acetic acid and 0.1 mM EDTA] at 80 V. The remaining 20 µl was purified using a DNA clean and concentrator kit before ligation

into the pGEM-T® vector (T-vector) using a ratio of 3:1 with 1 x ligation buffer and 1 U of T4 DNA ligase at 4ºC overnight.

The ligation mix was transformed into competent *E. coli* JM109 cells using the TransformAid™ bacterial transformation kit according to the manufacturer's guide. *E. coli* JM109 cells (50 µl) from each transformation were plated on pre-warmed 2x YT plates [1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl, 15 g/l bacteriological agar] containing 100 µg/ml ampicillin, X-gal (20 µg/ml) and IPTG (10 µg/ml) before overnight incubation at 37ºC. White colonies containing recombinant T-vector were selected for colony PCR using the respective CATL region primers described above and products analysed on a 1% (w/v) agarose gel in 1 x TAE buffer at 80 V. The positive clones were grown overnight in 5 ml of 2x YT medium before plasmid DNA extraction using the GeneJET™ plasmid miniprep kit according to the manufacturer's specifications. Recombinant T-vector clones were sequenced [Segoli sequencing unit, International Livestock Research Institute (ILRI), Nairobi, Kenya].

Recombinant *Tco*/*Tb*CATL T-vector plasmid DNA was restricted using XbaI and the resulting 320 bp (*Tco*CATL) and 300 bp (*Tb*CATL) products were subcloned into XbaI linearised p2T7^{Ti} RNAi vector (Fig. 2.1). Briefly, 50 µl of extracted plasmid DNA was restricted with XbaI to release T-vector cloned products and separated on a 1% (w/v) agarose gel in 1 x TAE buffer at 80 V. The 320 bp and 300 bp products were cut out of the gel and purified using the E.Z.N.A® gel extraction kit (PEQlab, Erlangen, Germany) according to the manufacturer's manual and eluted in 30 µl of elution buffer. The p2T7^{Ti} RNAi vector was also restricted with XbaI, purified and concentrated using the DNA clean and concentrator kit before treatment with 1 U of SAP for dephosphorylation to avoid the empty vector from re-ligating. The dephosphorylated vector and restricted fragments were then used in separate ligation reactions at a ratio of 1:3 (vector: insert) with 1 x ligation buffer and 1 U of T4 DNA ligase at 4ºC overnight before transforming into *E. coli* JM109 cells and plating on 2x YT plates containing 100 µg/ml ampicillin. Positive clones were identified using colony PCR with insert primers and confirmed with restriction using XbaI. Recombinant p2T7^{Ti}/*Tco*CATL (10 µg) was linearised with NotI in preparation for trypanosome transformation.

2.2.2.2 Trypanosome culture and transformation for *Tco*CATL RNAi

T. congolense TRUM183:29-13 procyclic parasites (Fig. 2.1) were propagated in complete MEM [MEM base powder, 25 mM HEPES, 26 mM $NaHCO_3$, 0.1 mM hypoxanthine, 20% (v/v) foetal calf serum, 4 mM haemin, 2 mM glutamine, 10 mM proline] supplemented with 2.5 µg/ml neomycin and 6.25 µg/ml hygromycin. *T. congolense* TRUM183:29-13 bloodstream form parasites isolated from mouse blood were maintained in complete IMEM [MEM base powder, 25 mM HEPES, 5.5 mM D-glucose, 1 mM sodium pyruvate, 0.04 mM adenosine, 0.1 mM hypoxanthine, 0.02 mM thymidine, 0.02 mM bathocuprone sulfate acid, 2 mM glutamine, 0.2 mM 2-mercaptoethanol, 0.5% (v/v) red blood cell lysate and 20% (v/v) goat serum]. *T. brucei* T7T-29 bloodstream form parasites were cultured in supplemented IMDM [IMDM base powder, 3.6 mM $NaHCO_3$, 1 mM hypoxanthine, 1 mM sodium pyruvate, 0.16 mM thymidine, 0.05 mM bathocuprone sulfate acid, 1.5 mM L-cysteine and 0.2 mM 2-mercaptoethanol, 10% (v/v) foetal calf serum] supplemented with 2.5 µg/ml neomycin and 6.25 µg/ml hygromycin.

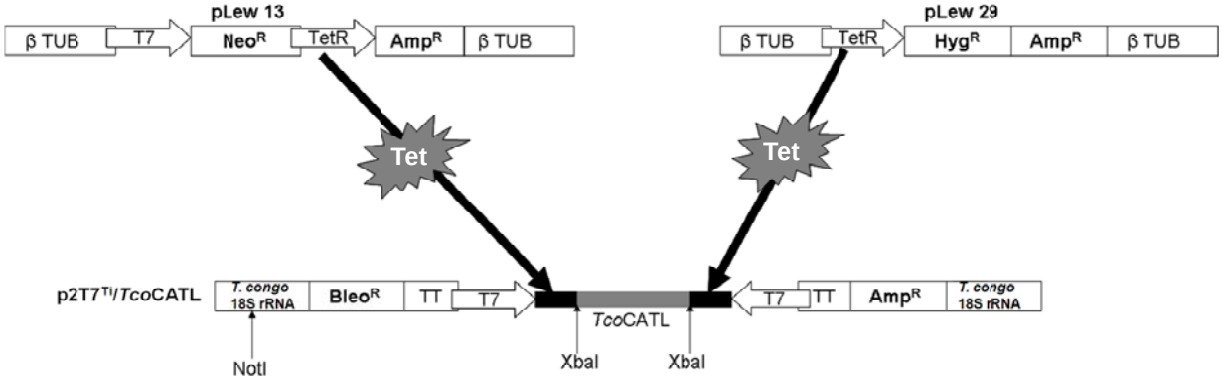


Figure 2.1 *Tco*CATL RNAi in *T. congolense*. *T. congolense* TRUM 183:29-13 parasites, previously transformed with pLew 29 and pLew 13 plasmids that code for a tetracycline repressor (Tet^R) which is able to suppress RNAi by binding to the tetracycline operator regions (▬), found on the $p2T7^{Ti}$ plasmid were used. Addition of tetracycline (Tet) blocks the repressor and allows transcription of the T7 polymerase T7 which produces dsRNA that is channelled into the endogenous RNAi pathway found in *T. congolense*. Other regions found on the $p2T7^{Ti}$ plasmid are TT sites (T7 transcription terminators), $Bleo^R$ (Bleomycin resistance), Amp^R (ampicillin resistance), XbaI sites for bacterial cloning, and a NotI site for integrating the $p2T7^{Ti}$ into the *T. congolense* 18S rRNA spacer region. pLew 13 and pLew 29 possess neomycin (Neo^R) and hygromycin (Hyg^R) resistance respectively. They also both possess β tubulin non coding regions for integration into the *T. congolense* genome. Adapted from Inoue (2002) and Wirtz et al. (1999).

Transformation of T. congolense and T. brucei parasites for RNAi was carried out as previously described (Coustou et al., 2010). Briefly, 1 x 10^7 parasites were pelleted by centrifugation (2000 *g*, 10 min, RT) and resuspended in 100 µl of Amaxa Basic Parasite Nucleofactor H solution 2. NotI linearised $p2T7^{Ti}$/*Tco*CATL (10 µg) was added and the

mixture electroporated using the Amaxa Nucelofaction® transformation system (program X-100). The parasites were immediately transferred to 10 ml of pre-warmed complete MEM medium and grown for 24 h at 27ºC in a 25 cm2 flask. The 10 ml culture was added to 90 ml of selection medium containing bleomycin (2 µg/ml), neomycin (2.5 µg/ml) and hygromycin (6.25 µg/ml). Two ml per well of the diluted culture was dispensed in 24 well plates. Wells containing bleomycin/neomycin/hygromycin resistant parasites were expanded to 10 ml and selected for genomic DNA isolation and PCR using primers specific to bleomycin [forward, (5'-ATGGCCAAGTTGACCAGTGCC-3') and reverse, (5'-TGCACGCAGTTGCCGGCCGGG-3')] to confirm integration of the p2T7^{Ti}/*Tco*CATL plasmid.

2.2.2.3 Assessment of infection and RNAi in mice using *Tco*CATL recombinant parasites

Transformed procyclic clones positive for the bleomycin-resistance gene were converted into bloodstream form parasites as previously described (Coustou *et al.*, 2010). Briefly, a 10 ml culture of procyclic parasites (10^7/ml) was pelleted by centrifugation (2000 *g*, 10 min, RT) and resuspended in incomplete MEM medium [MEM, Section 2.2.2 without foetal calf serum] for 2 h or until epimastigotes could be observed adhering to the side of the flask. The incomplete MEM was supplemented with 10% (v/v) foetal calf serum and half the medium changed every two days. This was maintained until the epimastigote cultures were confluent on the inside surface of the flask. The supernatant (10 ml) containing a mix of metacyclic and epimastigote trypanosomes was pelleted (2000 *g*, 10 min, RT), resuspended in 100 µl of complete MEM and injected intraperitoneally into NOD/SCID (NOD.Cg-Prkdescid Il2rgtm1 Wjl/Szj) mice. Mice were monitored daily for parasitaemia and once parasites were observed, blood was drawn using cardiac puncture and used to infect two groups of 5 Balb/C mice [previously immune-suppressed with cyclophosphamide (200 mg/kg)]. When parasitaemia in both groups of mice reached an average level of between 5 x 10^6 and 1 x 10^7, one group was used to induce RNAi *in vivo* by the addition of doxycycline hyclate (1 mg/ml) to their drinking water as previously described (Abdulla *et al.*, 2008). The mice were bled periodically from the tail to measure parasitaemia. An attempt to culture the bloodstream parasites *in vitro* in complete IMEM using bovine atrial endothelial (BAE) cells as feeder cells was made as previously described (Coustou *et al.*, 2010).

Blood samples were also collected and *Tco*CATL activity assayed by measuring the hydrolysis of Z-Phe-Arg-AMC. Briefly, a volume of mouse blood (determined by the level of parasitaemia in the different groups of mice) containing 1 x 10^5 parasites was diluted in 0.1%

(w/v) Brij-35 with 10 µg/ml SBTI, 1 µg/ml of pepstatin A, 1 mM AEBSF and 1 mM EDTA to make a total of 25 µl and incubated with 50 µl assay buffer [200 mM Tris-HCl buffer, pH 8, 10 mM DTT and 0.02% (w/v) NaN_3] for 10 min at 37°C. Aliquots were combined with substrate Z-Phe-Arg-AMC (20 µM) and the fluorescence read (excitation at 360 nm and emission at 460 nm) using a FLUOstar OPTIMA fluorescence microplate reader (BMG Labtech, Offenburg, Germany). For statistical analyses, values were expressed as means ± standard error of the mean (SEM). Significance levels were calculated by unpaired t tests and differences were considered significant at a *p value <0.001.

2.2.2.4 Identification of *Tco*CATL and *Tb*CATL 3' and 5' regions flanking the CATL tandem gene array

The *T. brucei* tandem gene array was identified in the GeneDB database (http://www.genedb.org/genedb/tbrucei/, accessed 10-10-2007) and used to design primers that amplified regions flanking the 5' and 3' regions of the 19 kb *TbCATL* gene array (Table 2.1) (see Appendix 1 for gene array). In *T. congolense,* selecting the flanking regions for *TcoCATL* was more difficult than in *T. brucei* because the *T. congolense* genome has not been fully sequenced. The gene that codes for *TcoCATL* is found in a multicopy array that is arranged in tandem repeats of more than 20 genes in units of ~2 kb when restricted with PstI (Fig. 2.2). A survey of *TcoCATL* genes in the GeneDB database (http://www.genedb.org/genedb/tcongolense/, accessed 10-10-2007) using the gene sequence for *TcoCATL* (congopain) [EMBL accession no. L25130, (Jaye *et al.*, 1994)] revealed two contig sequences (0001218 and 0002521) that contained several *TcoCATL* regions (see Appendix 2 for alignment). The 0001218 contig contained a ~ 6.9 kb PstI fragment on the 5' end of a *TcoCATL* gene (Fig. 2.2). This was taken as the 5' flanking region of the gene array and two primers spanning a 997 bp section of the 5' flanking region were designed (Table 2.1). In order to identify the 3' flanking region of the *TcoCATL* array (Fig. 2.2), a mix of cosmid vectors enriched for *TcoCATL* genes was restricted with PstI, separated on a 0.8% (w/v) agarose gel, transferred onto DNA binding nylon membrane using capillary action and cross-linked with UV before pre-hybridisation at 42 °C and subsequent hybridisation with biotin labelled probes generated from the *TcoCATL* ORF. Biontinylated probes were prepared by first perfoming a PCR using primers that amplified the pro-enzyme coding region at 1260 bp (Forward 5'-GCCGAATTCGCGTGCTCTGTTCCCGTGGCG-3' and Reverse 5'-CAGGCG GCCGCCTCGTGCCGCACGAGCCGAGC-3'). The resulting 1260 bp *Tco*CATL product was labelled using the Biotin DecaLabel™ DNA labeling kit according to the

manufacturer's instructions. Briefly, the PCR product was incubated with reaction buffer, biotin-labelled dNTPs from the kit and klenow fragment before incubation at 37°C overnight. The labelling mix was added to the nylon membrane, washed with stringency buffer [0.3 M NaCl, 30 mM sodium citrate, pH 7.0 containing 0.1% (v/v) SDS] and revealed using 0.06% (w/v) 4-chloro-1-naphthol, 0.0015% (v/v) H_2O_2 in Tris buffered saline (20 mM Tris–HCl, 200 mM NaCl, pH 7.4). A 2.3 kb region identified as the 3' flanking region of the gene array and which did not correspond to 2 kb or multiples of 2 kb was cloned into a PstI restricted T-vector and sequenced (Segoli sequencing unit, ILRI, Nairobi, Kenya). The resulting sequence was used to design primers that amplified a 600 bp region of the 3' flanking region (Table 2.1).

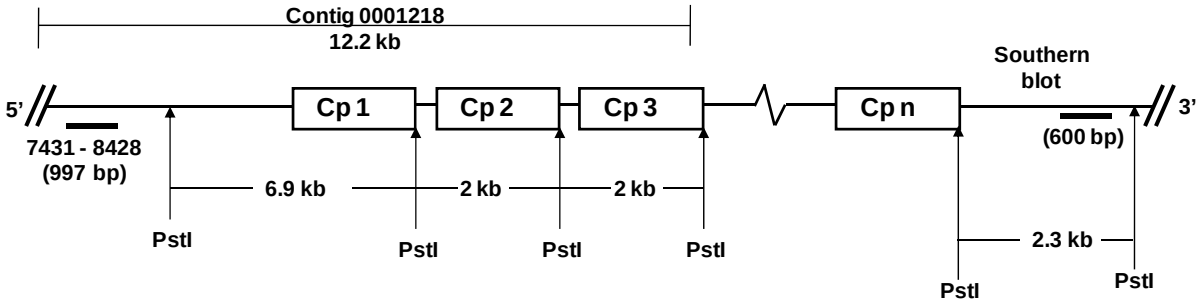


Figure 2.2 Arrangement of *Tco*CATL genes within the *TcoCATL* gene array. Restriction of the array using a PstI site located on the 3' end of each *TcoCATL* gene copy (indicated as Cp1, Cp2, Cp3...Cpn) gives a two kb fragment except for the two extreme ends of the array. The 5' end gives a fragment of approximately 6.9 kb in size (GeneDB), and the 3' end gives a fragment of approximately 2.3 kb (Southern blot on Section 2.3.4).

2.2.2.5 Cloning *Tco*CATL and *Tb*CATL flanking regions into knock-out vectors

T. congolense strain IL3000 and *T. brucei* 427 Lister genomic DNA were extracted from *in vitro* cultures as previously described (Medina-Acosta and Cross, 1993) and used as a template for PCR with primers shown in Table 2.1. Products from the 3' and 5' flank sets were cloned into a T-vector as described in Section 2.2.2 and recombinant clones restricted with appropriate enzymes for which restriction sites had been incorporated into the primers in preparation for subcloning into knock-out vectors (Fig. 2.3). The fragments generated after using the primers shown in Table 2.1 were serially cloned into two different knock-out vectors pGLneo and pGLbla (Fig. 2.3). Briefly, recombinant 5' and 3' flank T-vector plasmid DNA was restricted using appropriate restriction sites that had been incorporated into the primers (Table 2.1), resulting in 3' and 5' flanking regions for both *Tb*CATL and *Tco*CATL gene arrays. The resulting restriction products were serially subcloned into linearised pGLneo and pGLbla knock-out vectors (Fig. 2.3) as described in Section 2.2.2. Positive clones and the orientation of the flanks were confirmed using colony PCR with insert and vector primers

(Table 2.1). Ten µg from each of the four recombinant knock-out vectors was linearised with NotI in preparation for trypanosome transformation as described in Section 2.2.2.

Table 2.1 Primer sequences designed to amplify the 5' and 3' flanking regions from *T. congolense* strain IL3000 and *T. brucei* 427 Lister for cloning into the knock-out vectors pGLneo and pGLbla. Each primer insert had either a NotI, XbaI or ApaI restriction site (underlined) for incorporation into the pGLneo or pGLbla knock-out plasmids. Neomycin and blasticidin resistance gene primers were used to confirm orientation of the 3' flanking regions in the final recombinant plasmids. All primers were designed from database except for the *Tco*CATL 407 bp 3' pair which was designed from sequencing data derived from a Southern blot (Section 2.3.4).

Primer name and restriction site	Sequence (5'→3',restriction site underlined)
*Tco*CATL 997 bp 5' flank forward (NotI)	CGGCGGCCGCCACGCGGGTGAAGGCCACACAACAG
*Tco*CATL 997 bp 5' flank reverse (XbaI)	CGTCTAGATATTCTTATTTTTAGCTCGTTCGTG
*Tco*CATL 407 bp 3' flank forward (ApaI)	CGGGGCCCTTATTTTGTGTAAGCGTCTGGTGTA
*Tco*CATL 407 bp 3' flank reverse (ApaI)	CGGGGCCCTTGGGCAGGCGTAAAAAGGGTACCC
*Tb*CATL 511 bp 5' flank forward (NotI)	CGGCGGCCGCATCGAGCCTGTACCGTCGTGAAGGT
*Tb*CATL 511 bp 5' flank reverse (XbaI)	CGTCTAGATTTTTTGTGTACAGCAGGAGTGACGG
*Tb*CATL 390 bp 3' flank forward (ApaI)	CGGGGCCCCTATTTTCTTTTTTCCTTTCC
*Tb*CATL 390 bp 3' flank reverse (ApaI)	CGGGGCCCGATATCAAAAGTAAGGTAAGCGC
pGLbla forward primer (Bla Fw)	TCAACAGCATCCCCATCTCT
pGLneo forward primer (Neo Fw)	CGTGTTCCGGCTGTCAGCGC

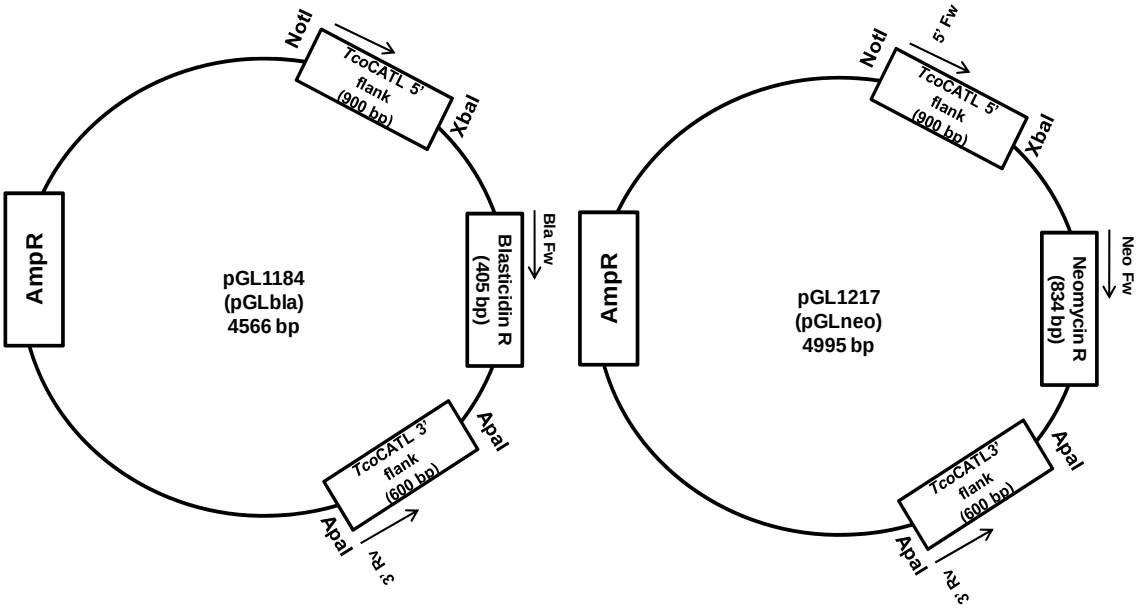


Figure 2.3 Knock-out plasmids for deletion of the *TcoCATL* array in *T. congolense* strain IL3000. pGLbla provides blasticidin resistance to transformed parasites, whereas pGLneo confers neomycin resistance. The *Tco*CATL 3' flank region in both plasmids uses an ApaI restriction site for sub-cloning and the 5'→3' orientation of the insert is confirmed by using the 3' reverse (Rv) insert primer with the Bla/Neo forward (Fw) vector primer. A single PCR product confirms that the orientation is correct. Inserts were also confirmed with PCR using 5' flank and 3' flank insert primers.

2.3 RESULTS

2.3.1 Cloning conserved *TcoCATL* and *TbCATL* gene regions into the p2T7^{Ti} RNAi vector

Genomic DNA was isolated from *T. brucei* 427 Lister and *T. congolense* IL3000 strains (Fig. 2.4, Panel A1, lanes 1 and 2) and used as a template for genomic DNA PCR with primers that had been designed to amplify conserved regions of the gene. A 300 bp product from *T. brucei* 427 Lister (*Tb*CATL) and a 320 bp product from *T. congolense* IL3000 (*Tco*CATL) were successfully amplified (Fig. 2.4, A2, lanes 1 and 2) and ligated with the pGEM-T® vector (T-vector). Blue white screening and colony PCR (Fig. 2.4, B) revealed five T-vector recombinant clones each for *Tco*CATL (lanes 1-5) and *Tb*CATL (lanes 6-10).

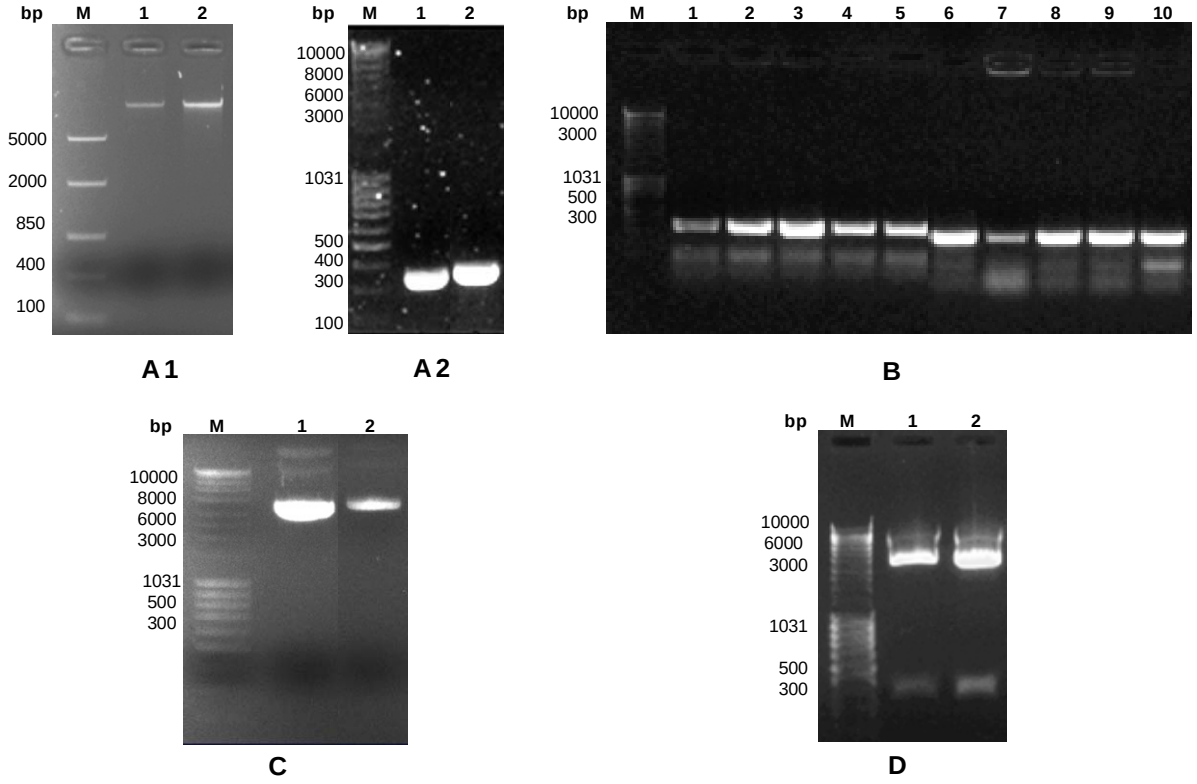


Figure 2.4 Amplification and T-vector cloning of *CATL* conserved gene regions from *TbCATL* and *TcoCATL*. M, MassRuler® DNA ladder mix / Middle Range® DNA ladder mix. **A1:** Genomic DNA isolated from **1**, *T. brucei* 427 Lister; **2**, *T. congolense* IL3000. **A2: 1**, 300 bp *TbCATL* PCR product using primers designed for *T. brucei* genomic DNA; **2**, 320 bp *TcoCATL* PCR product using primers designed for *T. congolense* genomic DNA. **B:** Colony PCR for **1-5**, *TbCATL* clones; **6-10**, *TcoCATL* clones. **C:** T-vector miniprep **1**, *TbCATL*; **2**, *TcoCATL*; **D:** XbaI confirmatory restriction **1**, 320 bp *TcoCATL*; **2**, 300 bp *TbCATL*. All samples were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml)**.**

Plasmid minipreparations from each construct were prepared for *TbCATL* and *TcoCATL* (Fig. 2.4, C, lanes 1 and 2) and confirmed using a restriction digest with XbaI giving a 320 bp for *TcoCATL* and a 300 bp product for *TbCATL* (Fig. 2.4, D, lanes 1 and 2).

Recombinant p2T7^{Ti}/BIP was prepared for subcloning by releasing BIP (1000 bp) using XbaI (Fig. 2.5, A, lanes 1 and 2). XbaI linearised p2T7^{Ti} vector (8000 bp) (Fig. 2.5, B1, lane 1 and C, lane 1) was ligated with 320 bp *Tco*CATL (Fig. 2.5, B2, lane 1) and 300 bp *Tb*CATL (Fig. 2.5, C, lane 2).

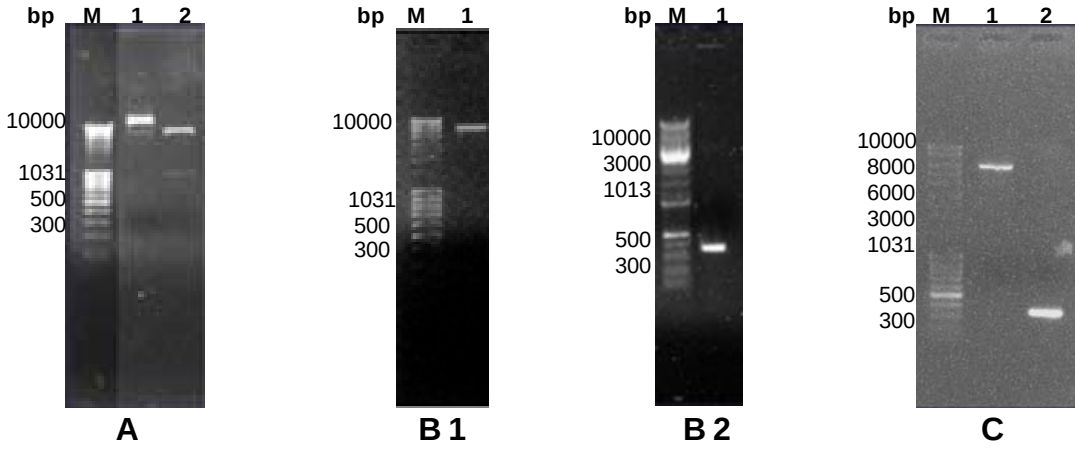


Figure 2.5 Subcloning of *CATL* conserved gene regions from *TbCATL* and *TcoCATL* into the p2T7^{Ti} RNAi vector. M, MassRuler® DNA ladder mix. **A: 1**, Miniprep of recombinant p2T7^{Ti}/BIP vector; **2**, XbaI restricted p2T7^{Ti} releasing ~ 1000 bp BIP. **B1: 1**, gel purified XbaI restricted p2T7^{Ti} vector. **B2: 1**, 320 bp gel purified *Tco*CATL fragment from T-vector. **C: 1**, gel purified XbaI restricted p2T7^{Ti} vector; **2**, 300 bp gel purified *Tb*CATL fragment from T-vector. Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml)**.**

Recombinant clones were selected for colony PCR using insert primers that amplified 320 bp products for recombinant p2T7^{Ti}/*Tco*CATL (Fig. 2.6, A, lanes 1 -10) and 300 bp products for recombinant p2T7^{Ti}/*Tb*CATL (Fig. 2.6, B, 1-3 and 5). Successful subcloning into the p2T7^{Ti} was confirmed using XbaI restriction for p2T7^{Ti}/*Tb*CATL showing a 300 bp band (Fig. 2.6, C1, lanes 1-3) and p2T7^{Ti}/*Tco*CATL showing a 320 bp band (Fig. 2.6, C2, lane 1). Recombinant p2T7^{Ti}/*Tco*CATL clones were selected for transformation into *T. congolense* TREU 183 29-13 procyclic parasites and restricted with NotI (Fig. 2.6, D, lanes 1-4). Recombinant p2T7^{Ti}/*Tb*CATL clones were also selected for transformation into *T. brucei* T7T-29 bloodstream form parasites and restricted with NotI (Fig. 2.6, D, lanes 5 - 8). Recombinant p2T7^{Ti}/*Tco*CATL was successfully transformed into procyclic *T. congolense* TREU 183 29-13

clones. Several rounds of transformations using p2T7^{Ti}/*Tb*CATL and *T. brucei* T7T- 29 bloodstream form parasites did not yield any viable clones (results not shown).

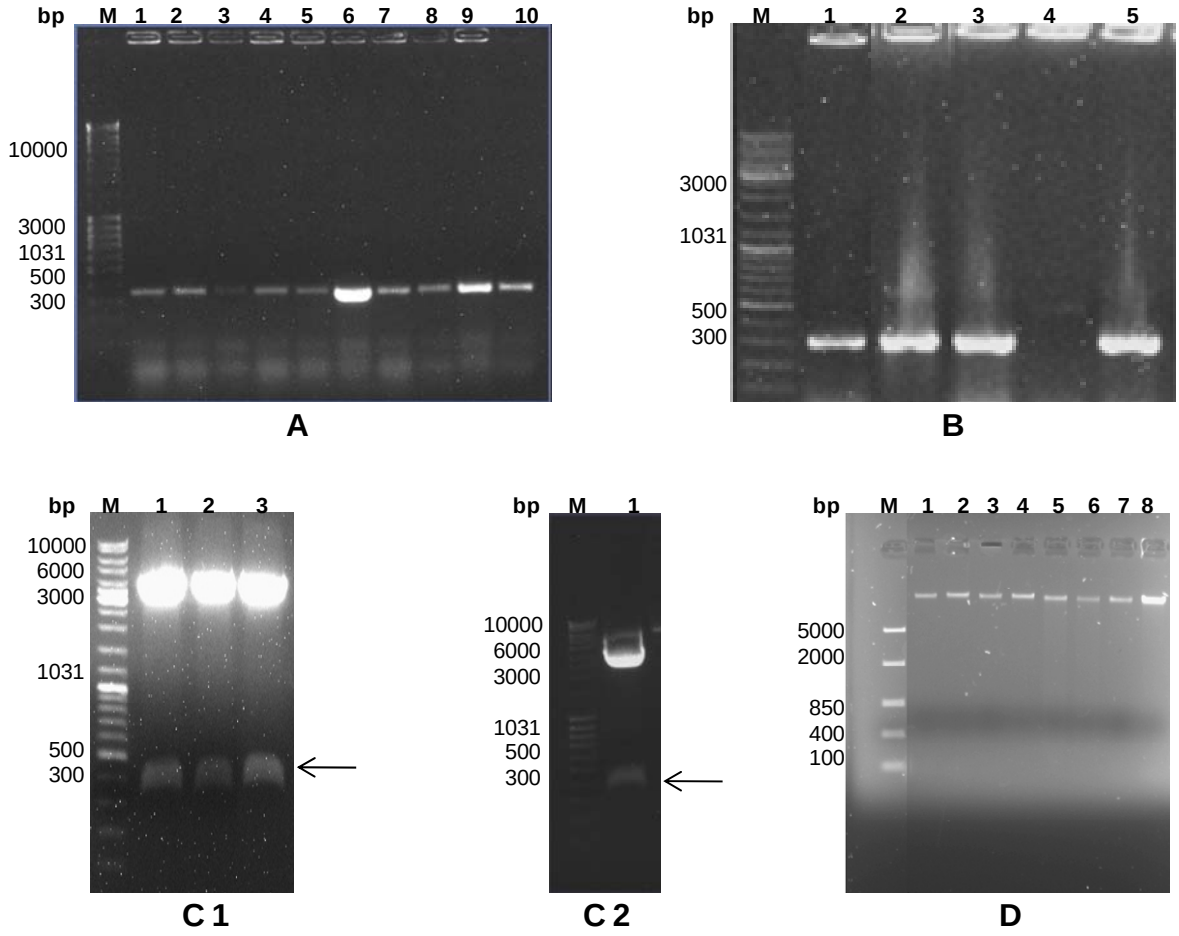


Figure 2.6 Subcloning of *CATL* conserved gene regions from *TbCATL* and *TcoCATL* into the p2T7^{Ti} RNAi vector. M, MassRuler® DNA ladder mix in all gels. **A: 1-10**, Positive p2T7^{Ti}/*Tco*CATL clones using 320 bp primers. **B: 1-3 & 5**, Positive p2T7^{Ti}/*Tb*CATL clones using 300 bp primers. **C1: 1-3**, p2T7^{Ti}/*Tco*CATL clones restricted with XbaI, releasing a 320 bp product (←). **C2: 1**, p2T7^{Ti}/*Tb*CATL clone restricted with XbaI, releasing a 300 bp product (←). **D: 1-4**, p2T7^{Ti}/*Tco*CATL clones restricted with NotI; **5-8**, p2T7^{Ti}/*Tb*CATL clones restricted with NotI in preparation for transformation. Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml)**.**

T. congolense TREU 183 29-13 p2T7^{Ti}/*Tco*CATL positive clones were expanded and genomic DNA isolated from cultures which was used in a PCR reaction with primers specific to bleomycin resistance. A 1000 bp product was detected in recombinant clones (Fig. 2.7, A, lanes 1 - 3) using the p2T7^{Ti} plasmid as a positive control (Fig. 2.7, A, lane 4) and genomic DNA from untransformed *T. congolense* TREU 183:29-13 parasites as a negative control (Fig. 2.7, B, lane 1).

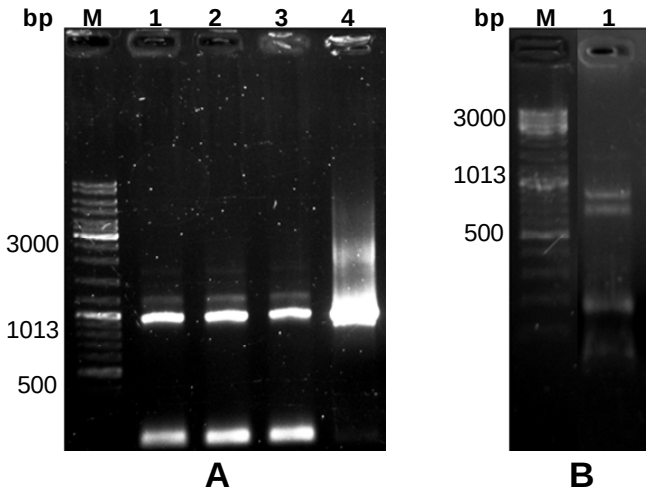
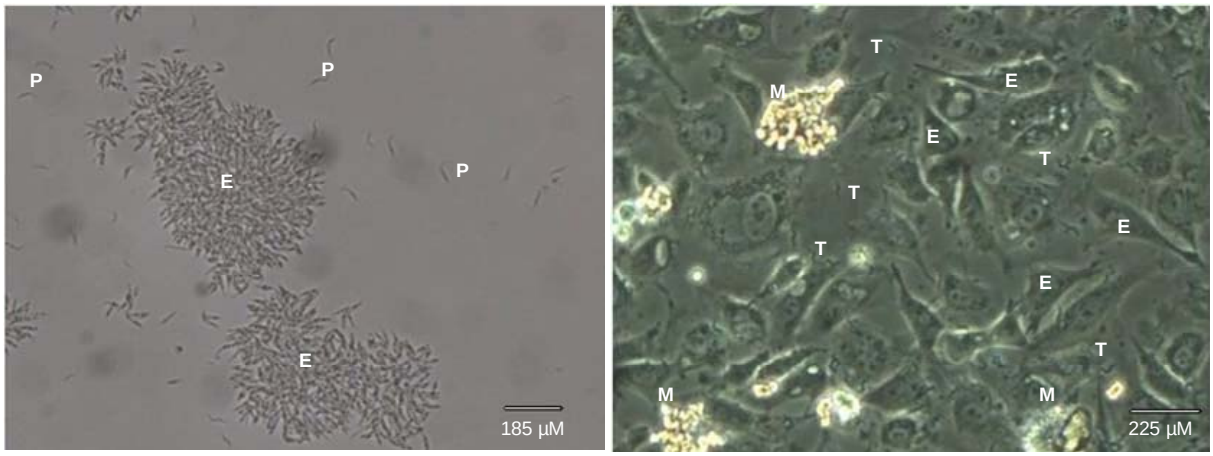


Figure 2.7 Confirmation of p2T7^{Ti} /*Tco*CATL insertion into *T. congolense* 29-13 strain IL3000 procyclic parasites. M, MassRuler® DNA ladder mix. **A: 1-3**, 1000 bp PCR product amplified from phleomycin resistant clones using primers specific to phleomycin resistance; **4.** p2T7^{Ti} plasmid used as a positive control. **B: 1**, untransformed *T. congolense* 29-1*3* strain IL3000 procyclic parasites used as a negative control. Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml).

2.3.2 Conversion of *T. congolense* TREU 183 29-13 p2T7^{Ti}/*Tco*CATL procyclic cultures into infective bloodstream forms

Recombinant *T. congolense* procyclic cultures were converted to bloodstream form parasites because *Tco*CATL expression during the procyclic stage of the trypanosome life cycle is very low (Mbawa *et al.*, 1991). Procyclic cultures were readily converted into epimastigote forms as evidenced by the presence of colonies of epimastigotes adhering to the tissue culture flask wall (Fig. 2.8, A). Metacyclic forms harvested from epimastigote cultures were then used to infect NOD mice and after several passages infected blood was used to infect Balb/C mice. Parasites were isolated from mouse blood and attempts to generate *in vitro* bloodstream form cultures did not result in long-term clones (Fig. 2.8, B). Bloodstream form parasites (T in Fig. 2.8, B) were co-cultured with bovine aortic endothelial cells (E in Fig. 2.8, B) in IMEM which acted as feeder cells as has been previously described (Coustou *et al.*, 2010). Cultures were passaged every 24 hours in an attempt to achieve stably dividing parasites. After 1 week of passages, most parasites would stop dividing and die (Fig. 2.8, B). The goat serum used to make complete IMEM was replaced with fresh goat serum as has been previously described (Coustou *et al.*, 2010), but this did not change the outcome of the experiment. After several attempts, it was decided to induce RNAi in the mouse model as has been successfully done before with *T. brucei* (Abdulla *et al.,* 2008).



A B

Figure 2.8 Conversion of recombinant p2T7^{Ti}/*Tco*CATL *T. congolense* procyclic cultures into bloodstream forms (BSF). A: Generation of epimastigotes from procyclics **P**, procyclic *T. congolense* parasites; **E**, epimastigote *T. congolense* parasite colonies adherent to the flask wall; **B**. Establishment of BSF cultures from infected mouse blood; **E**, Bovine aortic endothelial cells (BAE); **T**, *T. congolense* BSF; **M**, Residual mouse cells.

2.3.3 RNAi induction in mice infected with *T. congolense* TREU183:29-13 p2T7^{Ti}/*Tco*CATL parasites

When infection levels of *T. congolense* TREU 183 29-13 p2T7^{Ti}/*Tco*CATL bloodstream form parasites in mice were at comparable levels, RNAi was induced *in vivo* by the addition of doxycycline in the drinking water of one group of mice (+Tet). The control group was provided with regular water (-Tet). Hydrolysis of Z-Phe-Arg-AMC by plasma samples was used to track the activity of *Tco*CATL in +Tet mice and -Tet mice (Fig. 2.9, A). *Tco*CATL activity in both groups of mice declined as time progressed (Fig. 2.9, A). Z-Phe-Arg-AMC hydrolysing activity in +Tet mice reduced at a faster rate than in -Tet mice (Fig. 2.9, A). Parasitaemia was also measured in both groups of mice. The +Tet mice had a higher parasitaemia peak on day 8 but lower activity against Z-Phe-Arg-AMC than -Tet mice (Fig. 2.9, A and B). As the infection progressed in the two groups, parasitaemia in +Tet mice remained low but steadily increased in -Tet mice. Hydrolysis of Z-Phe-Arg-AMC in +Tet mice remained consistently lower than in -Tet mice. Parasitaemia on day 26 increased in both groups of mice and *Tco*CATL activity was almost as high as day one of induction. The experiment was stopped at this stage as the *T. congolense* TREU 183 29-13 p2T7^{Ti}/*Tco*CATL parasites caused a chronic infection in the mice. Parasitaemia between day 9 and 20 was lower in +Tet mice when compared to -Tet mice with the most significant difference on day 18 (Fig. 2.9, B). The most significant difference in the hydrolysis of Z-Phe-Arg-AMC between the two groups of test mice was seen on day 6 (Fig. 2.9, A).

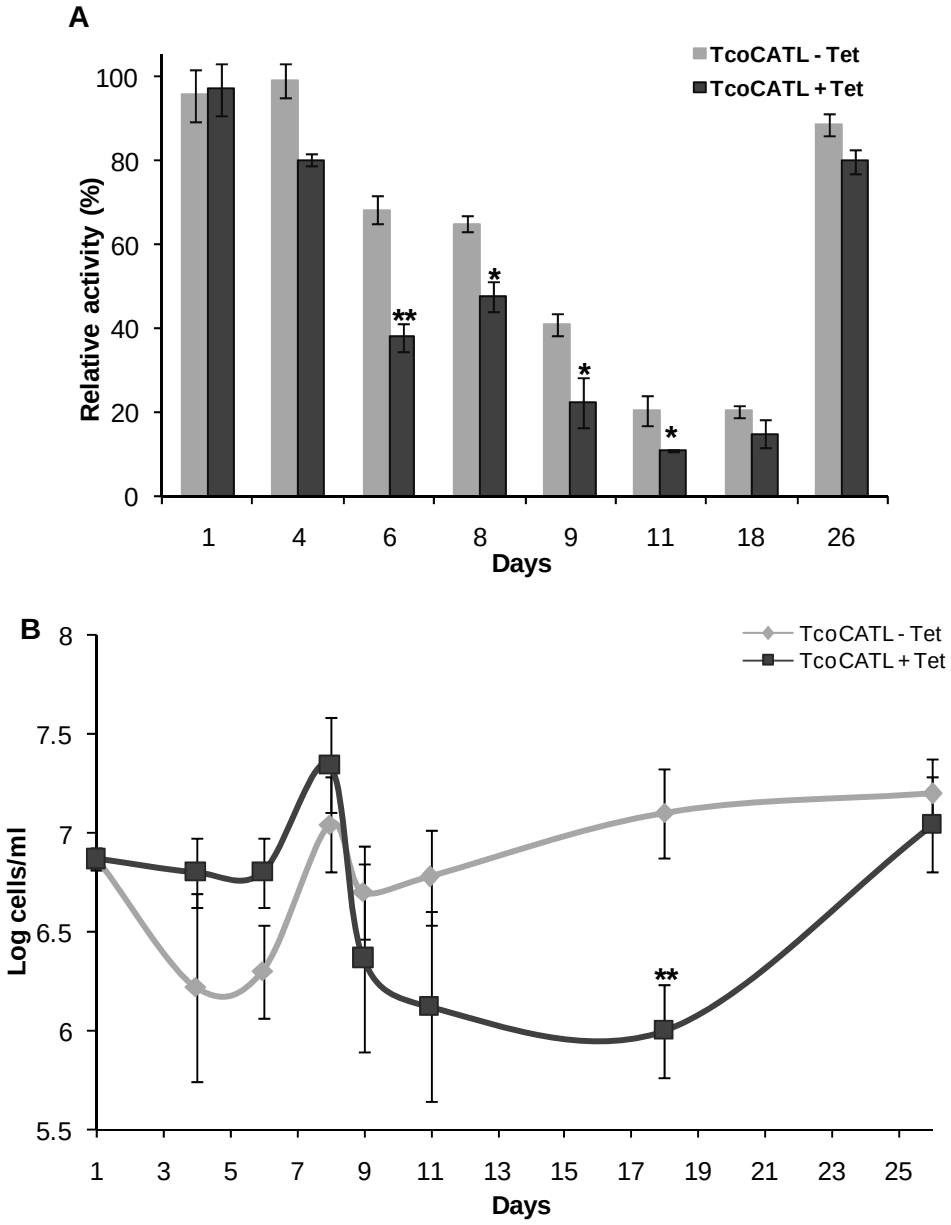


Figure 2.9 In vivo RNAi induction in mice infected with p2T7^{Ti}/*Tco*CATL *T. congolense* bloodstream form parasites. Two groups of 5 mice each were infected with *T. congolense* strain 29-13 parasites transformed with p2T7^{Ti}/*Tco*CATL. RNAi targeting *Tco*CATL was induced in one group by providing 1 mg/ml of doxycycline in the drinking water (+Tet) and the other group was used as a control (-Tet). **A:** A volume of blood from different days of infection containing 1 x 10^5 parasites from RNAi induced mice (+Tet) and non-induced mice (-Tet) was resuspended in 0.1% (w/v) Brij-35 with 10 µg/ml SBTI, 1 µg/ml of pepstatin A, 1 mM AEBSF and 1 mM EDTA to make a total of 25 µl and used in a fluorescence assay with Z-Phe-Arg-AMC. A mean of triplicate fluorescence readings was plotted relative to the highest reading as a percentage for each experiment ± SEM. **B:** Parasitaemia from RNAi induced mice (+Tet) was compared to non-induced control mice (-Tet). Each value is a mean of 5 mice ± SEM. *p< 0.05 and **p<0.001 in comparison to the non-induced control mice.

2.3.4 Identification of sequences flanking the 3' region of the *TcoCATL* gene array

Selecting the flanking regions for *TcoCATL* was difficult as the gene that codes for this particular protein exists as a multicopy array that is arranged in tandem repeats of more than 20 genes. A database search revealed a 6.9 kb fragment on the 5' flank of the array with one PstI site on the 5' end (Fig. 2.2). This fragment was used to design primers for the 5' region flanking the *TcoCATL* gene array and used to amplify a 997 bp product from *T. congolense* IL300 genomic DNA (Fig. 2.12, A1). The 3' region flanking the *TcoCATL* gene array could not be located in a database search, and required a different identification strategy. This involved restricting and probing recombinant cosmid vectors that have been enriched for *Tco*CATL genes (Boulangé *et al.,* unpublished results). The *Tco*CATL recombinant cosmid vectors were restricted with PstI which targets the 3' end of each *Tco*CATL gene in the array (Fig. 2.2)

Three recombinant cosmids enriched for *TcoCATL* genes were incompletely digested with PstI and separated on an agarose gel (Fig. 2.10, A) and a Southern blot performed (Fig. 2.10, B). Probing the Southern blot with biotin tagged full length *Tco*CATL revealed several bands that were 2 kb and multiples of 2 kb in size. A band at 2.3 kb was also identified (Fig. 2.10, B). The 2.3 kb band was cut out of the agarose gel (Fig. 2.10, C1, lane 2) and cloned into PstI restricted pGEM-T vector (Fig. 2.10, C1, lane 1). Recombinant 2.3 kb PstI fragment T-vector clones were identified using blue white screening and colony PCR with T7 and SP6 vector primers (Fig. 2.10, C2, lanes 1 and 2). A 1.9 kb insert previously cloned into the pGEM-T vector was used as a positive control (Fig. 2.10, C2, lane 3).

A plasmid preparation of one of the recombinant clones (Fig. 2.10, C3, lane 1), was confirmed using PstI restriction (Fig. 2.10, C3, lane 2) releasing empty 3 kb pGEMT-vector and the 2.3 kb PstI fragment. The two recombinant clones (Fig. 2.10, C2, lanes 1 and 2), referred to as pstfrag3 and pstfrag4 respectively, were sequenced and results revealed a region of ~ 490 bp that did not align to the intergenic region in the *Tco*CATL gene array (Fig. 2.11). This was expected as the both the 3' and 5' regions flanking the entire gene array are not expected to align with the *TcoCATL* intergenic region found within the array (as shown in Fig. 2.2). The region that did not align was identified as the 3' region flanking the *Tco*CATL gene array and used to design primers that amplified a 407 bp product (Fig. 2.12, B1).

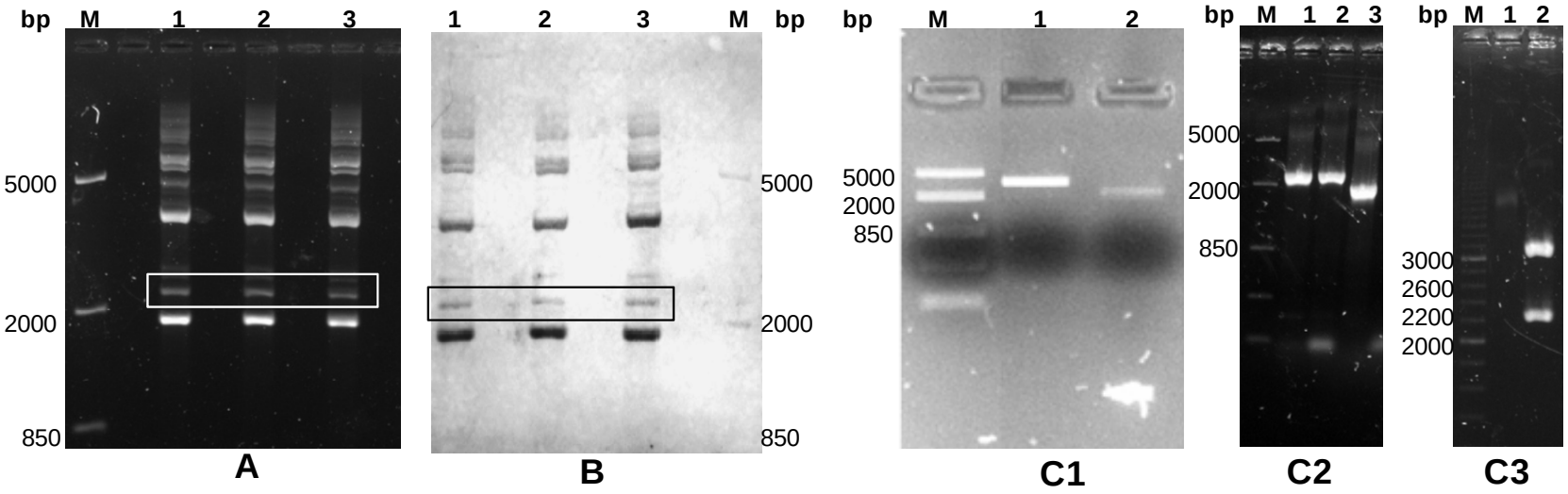


Figure 2.10 Identification of the 3' flanking region of the *Tco*CATL tandem gene array using Southern blotting. M, Middle Range® DNA ladder mix. **A: 1-3**, agarose (0.8%, w/v) gel electrophoresis of a mix of cosmid vectors enriched for *Tco*CATL genes restricted with PstI. Boxed area shows DNA that is not 2 kb or a multiple of 2 kb. **B: 1 - 3**, corresponding Southern blot of PstI restricted cosmid vectors enriched for *Tco*CATL genes developed with a biotinylated probe prepared from full length *Tco*CATL. Boxed area contains a 2.3 kb region identified as the 3' flank. **C1: 1**, PstI restricted gel purified pGEM T-vector prepared for ligation; **2**, gel purified 2.3 kb PstI restricted fragment identified using Southern blotting. **C2: 1-2**, Positive 2.3 kb PstI fragment T-vector clones identified by colony PCR using vector primers; **3**, 1.9 kb recombinant T-vector clone used as a positive control. **C3: 1**, Miniprep of positive 2.3 kb PstI T-vector clone; **2**, confirmatory restriction of positive 2.3 kb PstI T-vector clone using PstI. Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml).

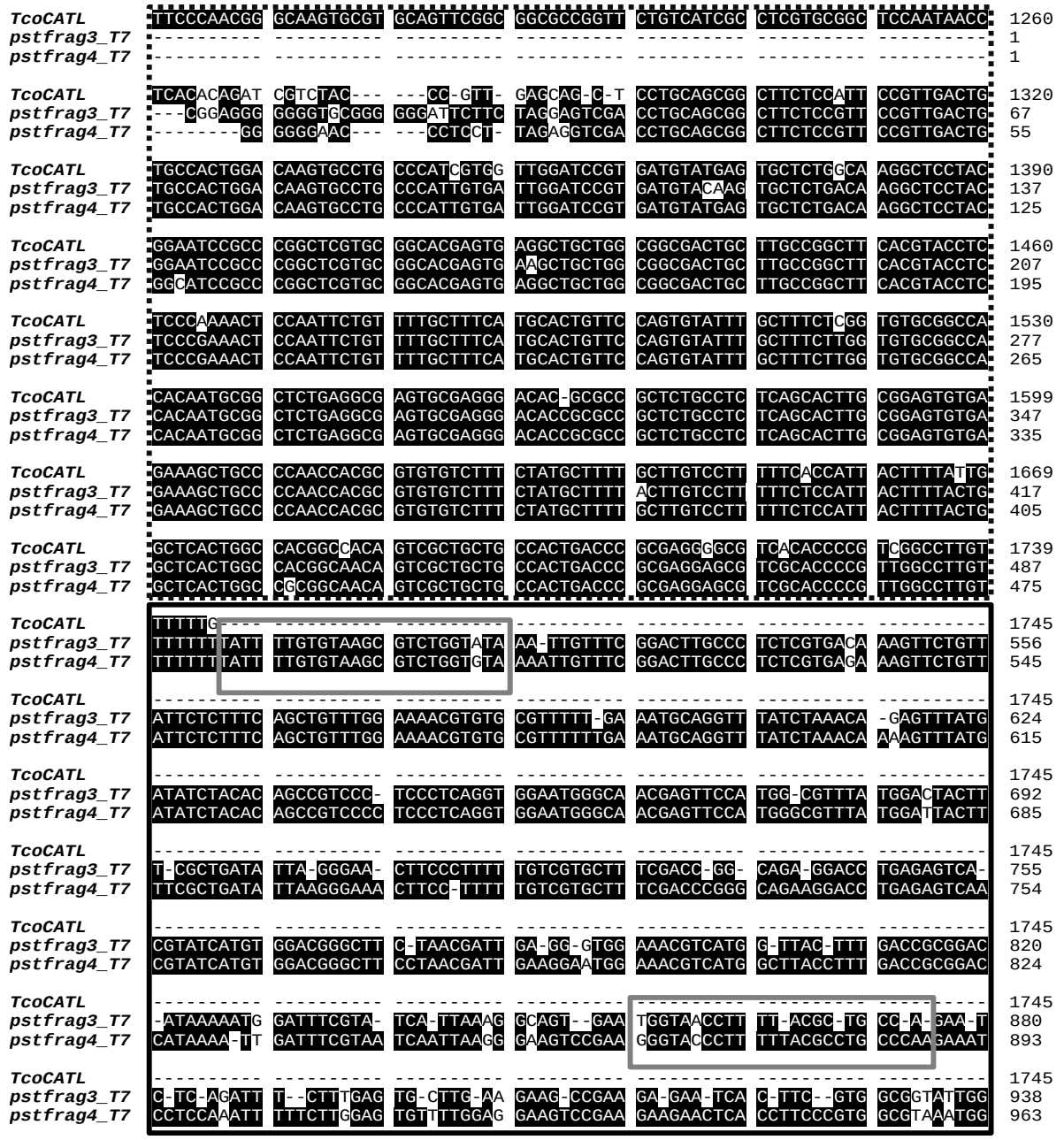


Figure 2.11 Sequence alignment of the two recombinant 2.3 kb PstI fragments pstfrag3 and pstfrag4 of the *Tco*CATL 3' flanking region in a pGEM T-vector. Sequencing was performed using T7 and SP6 primers that are specific for the pGEM T-vector. Sequences in the dashed box show the alignment between the extreme 3' end and the intergenic region of the *T. congolense* CATL array; sequences in the solid black box show the region downstream of the 3' end of the *Tco*CATL array, which align in two different clones of the 2.3 fragment in the pGEM T-vector; sequences in the grey boxes show regions that were used to design forward and reverse primers for cloning into knock-out vectors. Sequences were aligned using Clustal W.

2.3.5 Cloning the 3' and 5' flanking regions from the *TcoCATL* gene array into knock-out vectors pGLbla and pGLneo

The 997 bp 5'-*TcoCATL* flanking region PCR product, 5'-*Tco*CATL (Fig. 2.12, A1, lane 1) and the 407 bp 3'-*TcoCATL* flanking region PCR product 3'-*Tco*CATL (Fig.2, 212, B1, lane 1) were both cloned into a pGEM T-vector. Blue white screening and colony PCR were used to identify recombinant clones of 997 bp for 5'-*Tco*CATL (Fig. 2.12, A2, lanes 1 - 4) and 407 bp 3' *Tco*CATL (Fig. 2.12, B2, lanes 1 - 10). Recombinant clones were confirmed with restriction using NotI and XbaI for 5'-*TcoCATL* giving a 997 bp fragment (Fig. 2.12, A3, lane 1) and ApaI for 3'-*TcoCATL* giving a 407 bp fragment (Fig. 2.12, B3, lane 1), in addition to empty pGEM T-vector at 3 kb in both cases.

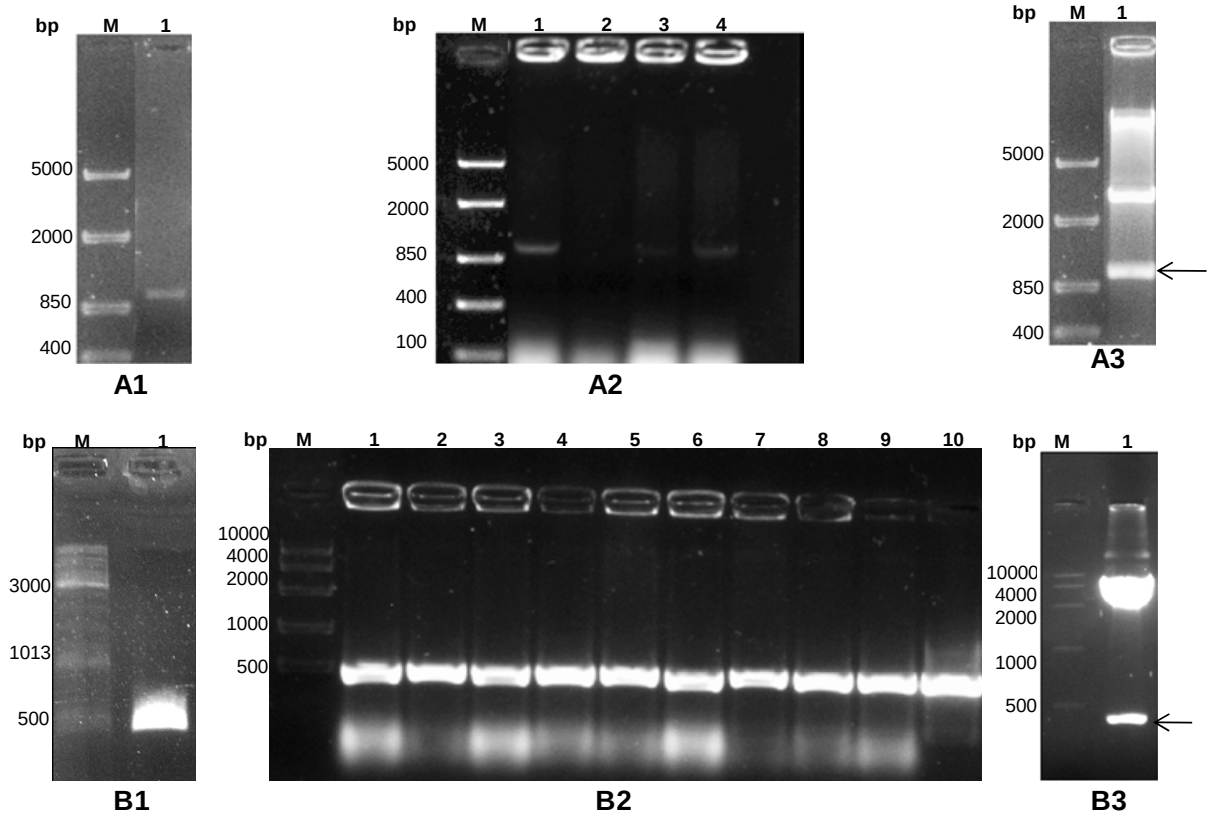


Figure 2.12 Amplification and T-vector cloning of 5' and 3' regions flanking the *Tco*CATL gene array. M, Middle Range® DNA ladder mix. **A1: 1**, 997 bp product flanking the 5' region of the *Tco*CATL gene array. **A2: 1 - 4**, colony PCR identifying positive 5' *Tco*CATL clones in T-vector. **A3: 1**, Confirmatory restriction of 997 bp 5' *Tco*CATL fragment in T-vector with NotI and XbaI (←). **B1: 1**, 407 bp product flanking the 3' region of the *Tco*CATL gene array. **B2: 1-10**, colony PCR identifying positive 3' *Tco*CATL clones in T-vector. **B3: 1**, Confirmatory restriction of 407 bp 3' *Tco*CATL fragment in T-vector with ApaI (←). Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml).

Vectors containing blasticidin resistance (pGLbla) and neomycin resistance (pGLneo) were prepared for ligation with 3'- and 5'- *Tco*CATL, or *Tb*CATL (Section 2.3.6) (Fig. 2.13, A, lanes 1 and 2). pGLbla (4566 bp) (Fig. 2.13, B, lane 1) and pGLneo (4995 bp) (Fig. 2.13, B, 2) were restricted with either NotI/XbaI or ApaI and before being serially ligated with the 997 bp 5'-*Tco*CATL and 407 bp 3'-*Tco*CATL fragments (Fig. 2.12, A3, lane 1 and B3 lane 1). Recombinant clones were confirmed using insert primers for 997 bp 5'-*Tco*CATL (Fig. 2.13, C lane 3 for 5'-*Tco*CATL/pGLbla, and D lane 3 for 5'-*Tco*CATL/pGLneo) and 407 bp 3'-*Tco*CATL (Fig. 2.13, C lane 2 for 3'-*Tco*CATL/pGLbla, and D lane 2 for 3'-*Tco*CATL/pGLneo).

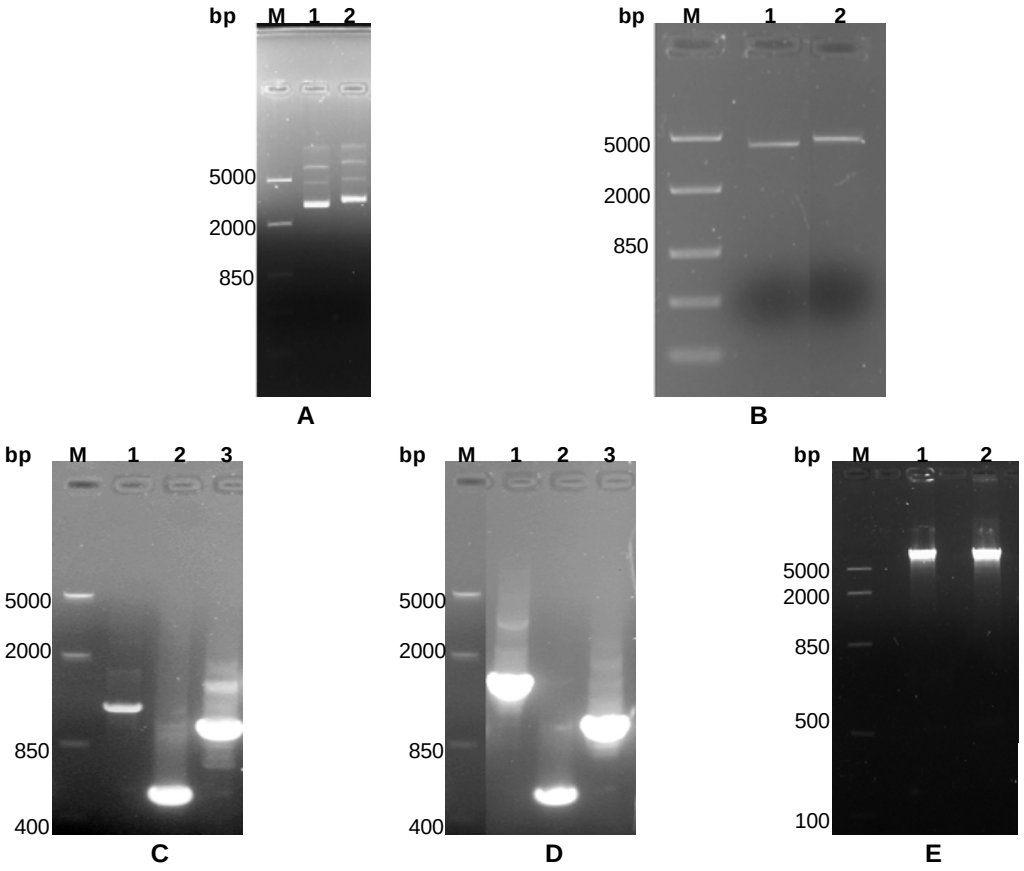


Figure 2.13 Subcloning of 5' and 3' regions flanking the *Tco*CATL gene array. M, Middle Range™ DNA ladder mix. **A:** Vector preparation - **1**, 4566 bp blasticidin resistance plasmid pGLbla (pGL1184); **2**, 4995 bp neomycin resistance plasmid pGLneo (pGL1217). **B:** Vector preparation - **1**, NotI/XbaI or ApaI restricted pGLbla; **2**, NotI/XbaI or ApaI restricted pGLneo. **C: 1**, 1074 bp product from recombinant pGLbla confirmed with PCR using blasticidin forward primer with insert 3' reverse primer; **2**, 407 bp product from recombinant pGLbla confirmed with PCR using 3' insert primers; **3**, 997 bp product from recombinant pGLbla confirmed with PCR using 5' insert primers. **D: 1**, 1213 bp product from recombinant pGLneo confirmed with PCR using blasticidin forward primer with insert 3' reverse primer; **2**, 407 bp product from recombinant pGLneo confirmed with PCR using 3' insert primers; **3**, 997 bp product from recombinant pGLneo confirmed with PCR using 5' insert primers. **E: 1**, 5970 bp NotI restricted recombinant *Tco*CATL/pGLbla; **2**, 6399 bp NotI restricted recombinant *Tco*CATL/pGLneo. Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml).

Orientation of the 3'-*Tco*CATL ApaI fragment in the 5' → 3' direction was confirmed using the 3'-*Tco*CATL insert reverse primer and the vector primer resulting in a 1074 bp product for 3'-*Tco*CATL/pGLbla (Fig. 2.13, C lane 1) and a 1213 bp product for 3' *Tco*CATL/pGLneo (Fig. 2.13, D, lane 1). Recombinant *Tco*CATL/pGLbla and *Tco*CATL/pGLneo clones were selected for transformation into *T. congolense* IL3000 procyclic parasites and restricted with NotI (Fig 2.13, E, lanes 1 and 2), resulting in a 5970 bp fragment for *Tco*CATL/pGLbla and a 6399 bp fragment for *Tco*CATL/pGLneo. Several rounds of transformations using *Tco*CATL/pGLbla and *Tco*CATL/pGLneo with *T. congolense* IL3000 procyclic parasites did not yield any viable clones (results not shown).

2.3.6 Cloning the 3' and 5' flanking regions from the *TbCATL* gene array into knock-out vectors pGLbla and pGLneo

Primers designed to amplify the 5'-*Tb*CATL flanking region (5'-*Tb*CATL) and the 3' flanking region (3'-*Tb*CATL) produced a 511 bp product for 5'-*Tb*CATL (Fig. 2.14, A1, lane 1) and a 390 bp product for 3'-*Tb*CATL (Fig. 2.14, B1, lane 1), which were each cloned into a T-vector. Blue white screening and colony PCR were used to identify recombinant clones for 5'-*Tb*CATL at 511 bp (Fig. 2.14, A2, lanes 1 - 4) and 3'-*Tb*CATL at 390 bp (Fig. 2.14, B2, lanes 1 - 3). Recombinant clones were confirmed using NotI and XbaI restriction for 5'-*Tco*CATL giving 511 bp (Fig. 2.14, A3, lanes 1 - 4) and ApaI restriction for 3'-*Tco*CATL at 390 bp (Fig. 2.14, B3, lanes 1 - 4).

pGLbla and pGLneo prepared for ligation with 3'- and 5'- *Tb*CATL (Section 2.3.5) and restricted with either NotI/XbaI or ApaI (Section 2.3.5), were serially ligated to the 511 bp 5'-*Tb*CATL fragment (Fig. 2.14, A3, lanes 1 - 4) and 390 bp 3'-*Tb*CATL (Fig. 2.14, B3, lanes 1 - 4) fragments. Recombinant clones were confirmed using insert primers for 511 bp 5'-*Tb*CATL (Fig. 2.15, A1, lane 1 for 5'-*Tb*CATL/pGLbla, and B, lane 3 for 5'-*Tb*CATL/pGLneo) and 390 bp 3'-*Tb*CATL (Fig. 2.15, A1, lane 2 for 3'-*Tb*CATL/pGLbla, and B, lane 2 for 3'-*Tb*CATL/pGLneo). Orientation of the 3'-*Tb*CATL ApaI fragment in the 5' → 3' direction was confirmed using the 3'-*Tb*CATL insert reverse primer and the vector primer resulting in a 1057 bp product for 3' *Tb*CATL/pGLbla (Fig. 2.15, A2, lane 1) and a 1196 bp product for 3'-*Tb*CATL/pGLneo (Fig. 2.15, B, lane 1). Recombinant *Tb*CATL/pGLbla (5467 bp) and *Tb*CATL/pGLneo (5896 bp) clones were selected for transformation into *T. brucei* 427 Lister bloodstream form parasites and restricted with NotI (Fig 2.15, C, lanes 1 and 2). Several

rounds of transformations using *Tco*CATL/pGLbla and *Tco*CATL/pGLneo with *T. brucei* 427 Lister bloodstream form parasites did not yield any viable clones (results not shown).

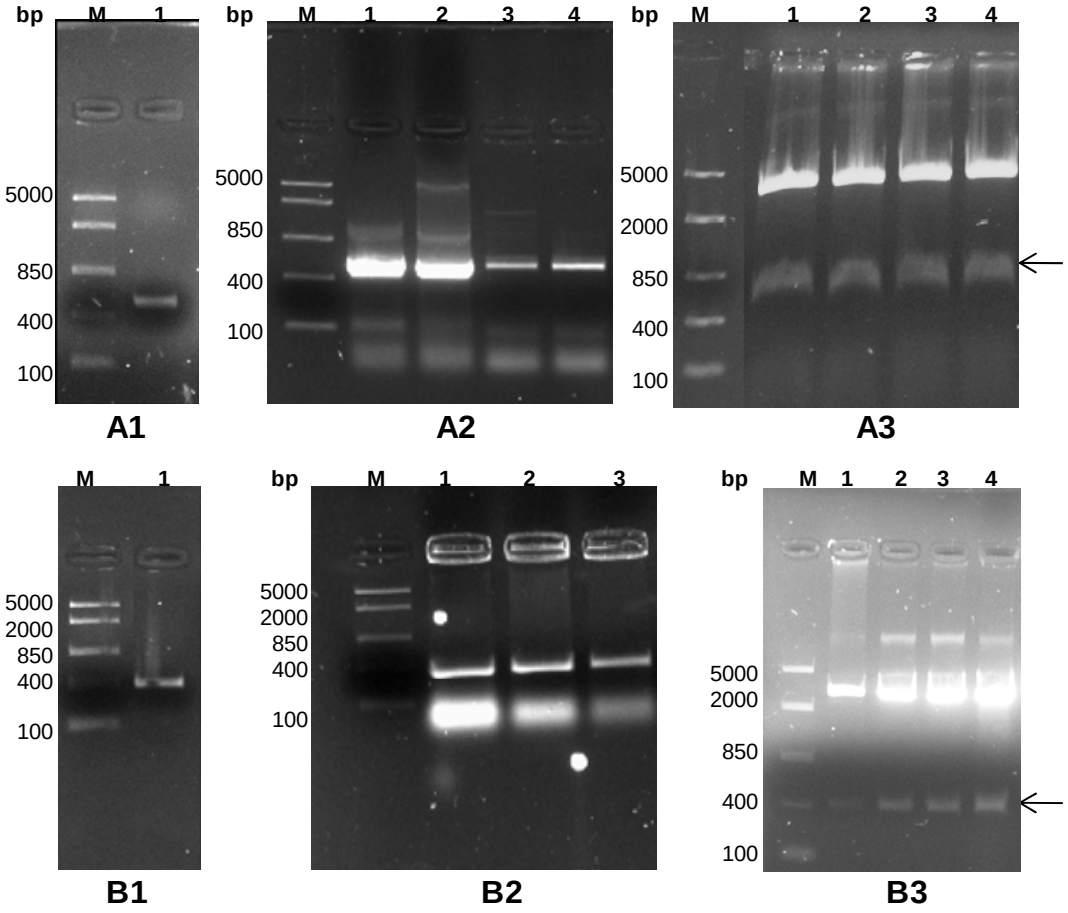


Figure 2.14 Amplification and T-vector cloning of 5' and 3' regions flanking the *TbCATL* gene array. M, Middle Range™ DNA ladder mix. **A1: 1**, 511 bp product flanking the 5' region of the *Tb*CATL gene array. **A2: 1-4**, colony PCR identifying positive clones from A1 in T-vector. **A3: 1 - 4**, confirmatory restriction of clone from A2 with NotI and XbaI releasing a 511 bp product as indicated by the arrow (←). **B1: 1**, 390 bp product flanking the 3' region of the *Tb*CATL gene array. **B2: 1-3**, colony PCR identifying positive clones from B1 in T-vector. **B3: 1**, confirmatory restriction of clone from B2 with ApaI releasing a 390 bp product as indicated by the arrow (←). Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml).

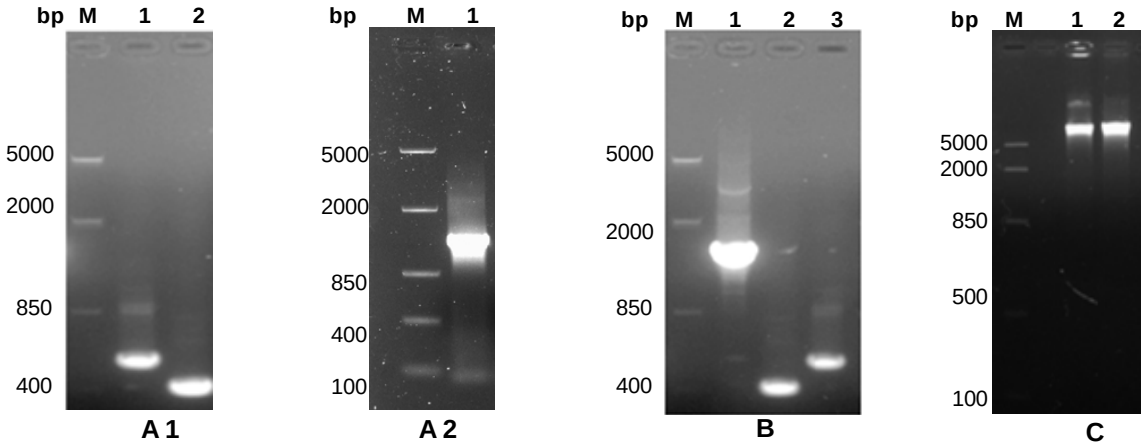


Figure 2.15 Subcloning of 5' and 3' regions flanking the *Tb*CATL gene array. M, Middle Range® DNA ladder mix. **A1: 1**, 511 bp product from recombinant pGLbla confirmed with PCR using 5' insert primers; **2**, 390 bp product from recombinant pGLbla confirmed with PCR using 3' insert primers. **A2: 1**, 1057 bp product from recombinant pGLbla confirmed with PCR using blasticidin forward primer with insert 3' reverse primer. **B: 1**, 1196 bp product from recombinant pGLneo confirmed with PCR using blasticidin forward primer with insert 3' reverse primer; **2**, 390 bp product from recombinant pGLneo confirmed with PCR using 3' insert primers; **3**, 511 bp product from recombinant pGLneo confirmed with PCR using 5' insert primers. **C: 1**, 5467 bp NotI restricted recombinant *Tb*CATL/pGLbla; **2**, 5896 bp NotI restricted recombinant *Tb*CATL/pGLneo. Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml).

2.4 DISCUSSION

Cysteine peptidases in Trypanosomes have been implicated in evading the host immune system, as well as cell and tissue invasion (Sajid and McKerrow, 2002). *Tco*CATL, the major cysteine peptidase in *T. congolense*, has been identified as an antigen during infection (Authié *et al.*, 1992) and is one of the parasite products that is responsible for host pathologies associated with the disease (Authié *et al.*, 2001). Experiments targeting the *T. brucei* cysteine peptidase brucipain (*Tb*CATL), for RNAi revealed that mice infected with parasites that have down-regulated levels of *Tb*CATL survived longer than mice infected with parasites that have regular levels of the peptidase (Abdulla *et al.*, 2008). The same study demonstrated a role for *Tb*CATL in helping *T. brucei* parasites cross the blood-brain barrier in an *in vitro* model (Abdulla *et al.*, 2008). *Tb*CATL targeted for RNAi as a control in the present study did not produce any transformed *T. brucei* T7T- 29 bloodstream parasites. Constructs targeting *Tco*CATL for RNAi on the other hand, were stably transformed into *T. congolense* TREU 183 29-13 procyclic cultures.

*Tco*CATL is only expressed during the bloodstream stage of *T. congolense* parasites, (Boulangé *et al.*, 2001), and it was therefore necessary to transform recombinant procyclic

cultures into bloodstream forms. This was successfully achieved using an approach that uses a transformation medium that converts *T. congolense* procyclic parasites into metacyclics (Coustou *et al.*, 2010). The metacyclics were then used to infect SCID mice and after several passages were used to infect BALB/c mice. Mice infected with parasites and targeted for *Tco*CATL RNAi had lower numbers of parasites than mice that were non-induced for RNAi. This observation was consistent with experiments carried out for *Tb*CATL RNAi (Abdulla *et al.*, 2008). *Tco*CATL activity was assessed under RNAi induction using fluorescence readings generated by the hydrolysis of Z-Phe-Arg-AMC, the preferred substrate for cysteine peptidase activity in trypanosomes (Mbawa *et al.*, 1991). Enzyme activity in mice where RNAi was induced with doxycycline was consistently lower compared to mice not supplied with doxycycline. This result indicated that the reduction in parasitaemia was as a direct result of the reduction of *Tco*CATL levels in RNAi induced parasites. It was noted that levels of activity in non-induced mice also declined. This was not surprising as the p2T7^{Ti} vector used for RNAi has leaky T7 promoters that are still able to transcribe the double stranded silencing RNA as has been previously described (Inoue *et al.*, 2002). This problem could have been alleviated by using a more stable RNAi vector such as pZJM used in RNAi experiments targeting *Tb*CATL for RNAi (Abdulla *et al.*, 2008).

In addition, it was also noted that parasites escape downregulation between days 18-25 (Fig. 2.9, panel B). This phenomenon has been described before where an essential genes targeted for RNAi can develop escape mutations as in the case of trypanothione in *T. brucei* (Ariyanayagam *et al.*, 2005). Downregulation of trypanothione was associated with several fold increases in free glutathione and a decrease in growth rate, but parasites targeted for RNAi reverted to normal glutathione levels and regular growth rates after two weeks. This was attributed to point mutations within the gene coding for trypanothione, thus enabling parasites to escape downregulation (Ariyanayagam *et al.*, 2005). Considering that *Tco*CATL is a gene coded for within an array of closely related variants (Pillay *et al.*, 2010), it is quite possible that related *Tco*CATL variants are able to become active in the absence of the major cysteine peptidase. This anomally can be avoided by targeting multiple variants for RNAi when downregulating the activity of *Tco*CATL, or by knocking out the entire gene array as was attempted in this study. Survival data of mice could not be obtained as the *T. congolense* TREU 183 strain used for this experiment causes an infection with low virulence that does not kill the mice (Coustou *et al.*, 2010).

Following successful down-regulation of *Tco*CATL using *in vivo* RNAi, an attempt was made to delete the *Tb*CATL gene array in a *T. brucei* 427 Lister strain and the *Tco*CATL gene array in a *T. congolense* IL3000 strain. The 5' and 3' regions flanking the *Tb*CATL gene array were readily identified on the GeneDB database where the full sequence of the *T. brucei* genome has been deposited (Berriman *et al.*, 2005). The 5' region flanking the *Tco*CATL gene array was also identified on the GeneDB database using contig 0001218 that contained a 6.9 kb PstI fragment bearing a *Tco*CATL gene copy on the 3' end of the contig. The 3' region flanking the *Tco*CATL gene array was unfortunately not available on the incomplete database and could only be identified using Southern blotting. Flanking regions for both gene arrays in *T. brucei* and *T. congolense* were successfully amplified and cloned into knock-out vectors. Attempts at transforming *T. congolense* and *T. brucei* parasites were, however, unsuccessful. This failure could be attributed to the size of the gene arrays in both parasites and several rounds of transformation would be required to get a successful clone. Because not all of the genes in the array code for active enzyme (Pillay *et al.*, 2010), an alternative to knocking out all the genes in the array would be to delete smaller sections and targeting copies of the gene that code for the mature active enzyme.

Cathepsin B (CATB) peptidases in trypanosomes belong to the same clan as CATL enzymes used in this study and are also inhibited by the diazomethane inhibitor Z-Phe-Ala-CHN_2 (Mackey *et al.*, 2004). *In vitro* and *in vivo* treatment of parasites with Z-Phe-Ala-CHN_2 was used to demonstrate the importance of cysteine peptidases for parasite survival (Scory *et al.*, 1999). Targeting *Tb*CATB for RNAi rather than *Tb*CATL *in vitro* validated the peptidase as a chemotherapeutic target resulting in endosome enlargement, defective cytokinesis and death of cultured parasites (Mackey *et al.*, 2004). *Tb*CATB is also necessary for degrading host proteins (Mackey *et al.*, 2004). Targeting *Tb*CATB for RNAi *in vivo* has a stronger phenotype than RNAi for *Tb*CATL, and mice infected with parasites induced for *Tb*CATB knockdown are able to survive and clear the infection (Abdulla *et al.*, 2008). With RNAi absent in Leishmania, a closely related kinetoplastid, knockout experiments of the *L. mexicana* cysteine peptidase gene array (*Lm*CPb) reduced macrophage infectivity by 80% indicating that *Lm*CPb is important for virulence (Mottram *et al.*, 1996).

In conclusion, this work was able to confirm that *Tco*CATL is an essential peptidase in *T. congolense* parasites during infection by inducing RNAi *in vivo* in mice and indicated that reduction in expression of the gene resulted in lower numbers of parasites when compared to

mice infected with non-induced parasites. In *T. brucei*, targeting *Tb*CATL negatively affects the ability of parasites to traverse the blood-brain barrier (Abdulla *et al.*, 2008), thus reducing their virulence. Mice infected with parasites induced for *Tb*CATL RNAi are therefore able to survive longer when compared to mice infected with parasites not induced for *Tb*CATL RNAi (Abdulla *et al.*, 2008). *T. congolense*, when compared to *T. brucei*, does not cross the blood-brain barrier or traverse blood endothelial cells into tissue and is thus restricted to the vascular system (Ojok *et al.*, 2002). Considering that *Tco*CATL is the major cysteine peptidase expressed during the bloodstream stage of the life cycle (Mbawa *et al.*, 1992), it is very probable that it plays multiple roles and down-regulating its expression results in a reduction of parasite viability as shown here. To further develop this work, it would be ideal to attempt the infection of cattle or larger ruminants with parasites induced for *Tco*CATL RNAi to observe if this would affect the progression of the disease in a natural setting. A reduced amount of circulating *Tco*CATL may result in reduced pathology associated with invariant antigens released by infecting parasites as observed in trypanotolerant cattle. This could potentially lead to the development of a vaccine or drug that targets *Tco*CATL, and may be an important step in controlling the disease in Africa.

It is difficult to generate *Tco*CATL null mutants for the sole reason that the gene exists in multiple copies. Fortunately this is not an obstacle in the case of OPB, another parasite peptidase that has been implicated as a pathogenic factor in trypanosomosis. *Tco*OPB and *Tb*OPB occur as a single copy gene in the respective trypanosomes. Studies that employ gene disruption techniques targeting OPB are described in chapters 3 and 4.

CHAPTER 3

IN VITRO DOWN REGULATION OF OLIGOPEPTIDASE B IN *TRYPANOSOMA CONGOLENSE* AND *TRYPANOSOMA BRUCEI* USING RNA INTERFERENCE

The African trypanosome species *T. congolense*, *T. vivax* and *T. b. brucei* are responsible for trypanosomosis in cattle on the African continent with a 100-fold higher prevalence of the disease in livestock than in humans (Baral, 2010). The main immune response by the host towards infecting bloodstream form parasites is the production of antibodies against the variable glycoprotein coat (VSG) found on the surface of the parasite (Taylor and Rudenko, 2006). Switching of VSG expression in a fraction of the infecting parasites occurs at genomic active expression sites as the host produces antibodies against the predominant VSG present in most of the circulating trypanosomes (Taylor and Rudenko, 2006). This ensures parasite survival and causes a new wave of parasitaemia that prolongs the disease (Pays *et al.*, 2004). Trypanosomes are also able to evade the immune system by endocytosis of VSG-antibody complexes, thereby protecting them from complement-mediated killing (Field *et al.*, 2009). Due to antigenic variation and immune evasion, it has been difficult to develop an effective vaccine that gives full protection. A more attractive option is to develop vaccines that target invariant antigens that do not change over the course of an infection (Antoine-Moussiaux *et al.*, 2009). When parasites undergo antibody mediated lysis, they release various compounds that could be responsible for pathogenesis in the mammalian host (Tizard *et al.*, 1978). Trypanosome peptidases released during lysis present attractive vaccine targets as they are invariant and remain active in the host bloodstream (Tizard *et al.*, 1978; Antoine-Moussiaux *et al.*, 2009). This has led to exploring the development of vaccines that target products released during an infection in a strategy referred to as an ‘anti-disease’ approach to reduce the symptoms of the disease rather than killing the parasite (Authié *et al.*, 2001; Schofield, 2007; Baral, 2010).

Several pathogenic trypanosomal products such as cysteine peptidases (Authié *et al.*, 1992; Boulangé *et al.*, 2001; Pillay *et al.*, 2010), paraglutamyl peptidases (Morty *et al.*, 2006), aminopeptidases (Pillay, 2011) and oligopeptidases (Morty *et al.*, 1999b; Morty *et al.*, 2005a; Coetzer *et al.*, 2008; Bastos *et al.*, 2010) have been characterised following recombinant expression. The definitive *in vitro* functions of these peptidases have not been extensively studied and RNA interference (RNAi) provides a viable option to elucidate their functions in trypanosomes. RNAi is described as the targeted degradation of specific RNA transcripts by

homologous short interfering RNAs (siRNA) generated from longer exogenous RNA introduced into the parasite (Ullu *et al.*, 2004; Balaña-Fouce and Reguera, 2007). Cysteine peptidases in *T. brucei* have been targeted for RNAi *in vitro* where the reduction of the cathepsin B-like peptidase (*Tb*CATB) led to parasites having an enlarged endosome, impaired parasite division and ultimately death (Mackey *et al.*, 2004). *T. congolense* parasites targeted for aminopeptidase RNAi *in vitro* displayed a growth defect in that they grew to a lower density when compared to parasites not targeted for aminopeptidase degradation (Pillay, 2011). Targeting the major surface peptidase in *T. brucei* procyclic parasites for RNAi led to a marked decrease in the release of transgenic variant surface glycoprotein 117 (VSG117) (LaCount *et al.*, 2003). *T. brucei* major surface peptidase mediated surface protease activity in *T. brucei* and inhibition of its activity with peptidomimetic collagenase inhibitors was toxic during the bloodstream stage of the life cycle (Bangs *et al.*, 2001). Metacaspases are cysteine peptidases found in trypanosomes but absent in the mammalian host (Berg *et al.*, 2010). This makes them attractive targets when designing drugs to combat trypanosomosis. *T. brucei* expresses five different metacaspases and experiments that target metacaspase 4 for RNAi and gene knockout showed that the peptidase is necessary for cell cycle progression and virulence during infection in mice (Proto *et al.*, 2011). It has also been shown that metacaspase 3 is responsible for processing metacaspase 4 once it is released by the parasite in a proteolytic cascade that generates a virulence factor necessary for establishing infections in the mammalian host (Proto *et al.*, 2011).

Oligopeptidase B (OPB) is a serine peptidase that has been identified as a possible pathogenic and virulence factor during infections caused by *T. brucei* (Morty *et al.*, 1999b), *T. congolense* (Morty *et al.*, 1999a), *T. cruzi* (Burleigh *et al.,* 1997) and *T. vivax* (Coetzer *et al.*, 2008). In rats infected with trypanosomes, OPB has been shown to remain active in the host bloodstream where it has been implicated in the hydrolysis of peptide hormones such as atrial natriuretic factor (ANF) (Morty *et al.*, 2005a) and gonadotropic releasing hormones (GnRH) (Tetaert *et al.*, 1993). In order to study what role OPB plays in *T. congolense* (*Tco*OPB) and *T. brucei* (*Tb*OPB) parasites *in vitro*, OPB knock-down parasites were generated using a tetracycline inducible RNAi vector (Inoue *et al.*, 2002). RNAi was monitored using anti-*Tco*OPB antibodies raised in chickens, and by analysing the reduction of RNA with reverse transcription PCR (RT-PCR) and quantitative PCR (RT-qPCR).

3.2 MATERIALS AND METHODS

3.2.1 Materials

Molecular biology: XbaI [for endonuclease nomenclature see, Roberts *et al.* (2003)], shrimp alkaline phosphatase (SAP), T4 DNA ligase, 10 mM dNTP mix, X-gal, IPTG, MassRuler® DNA ladder mix, Middle range® DNA ladder, GeneJET® plasmid miniprep kit, Transform Aid® bacterial transformation kit and RNase free DNAse were obtained from Fermentas (Vilnius, Lithuania). pGEM-T® vector was obtained from Promega (Madison, WI, USA). DNA clean and concentrator kit from ZymoResearch (Orange, CA, USA) and E.Z.N.A® gel extraction kit from PEQlab (Erlangen, Germany). *Escherichia coli* JM 109 and BL21 DE3 cells were purchased from New England Biolabs (Ipswich, MA, USA). FIREpol® Taq polymerase, 10 x PCR reaction buffer and 25 mM $MgCl_2$ were acquired from Solis Biodyne (Tartu, Estonia). Molecular biology grade agarose was purchased from Conda laboratories (Madrid, Spain), Ampicillin sodium salt from USB Corporation (Cleveland, OH, USA), bacteriological agar and tryptone from Merck (Darmstadt, Germany) and kanamycin sulfate was obtained from Fluka (Buchs, Switzerland). The yeast extract was purchased from Sigma (St. Louis, MO, USA). The recombinant $p2T7^{Ti}$/BIP RNAi plasmid was a kind gift from Professor Théo Baltz (University of Victor-Segalen, Bordeaux 2, France). Full length oligopeptidase B from *T. congolense* in pET28a was sourced from work previously done by Laura Huson and Lorelle Bizaaré in our laboratory (Huson, 2006; Bizaaré, 2008).

Recombinant *Tco*OPB purification and peptide hydrolysis: Imidazole and His-select® nickel affinity column was purchased from Sigma (St. Louis, MO, USA). Amicon Centriprep® concentration filters were obtained from Millipore (Bedford, MA, USA) and the 2 mg albumin standard was obtained from Thermo Fisher scientific (Rockford, IL, USA). The peptide substrate benzyloxycarbonyl (Z)-Arg-Arg-7-amino-4-methylcoumarin (AMC) was obtained from Bachem (Torrance, CA, USA).

Chicken IgY preparation and ELISA: Freund's complete and incomplete adjuvants, rabbit anti-IgY coupled to HRPO and bovine serum albumin (BSA) were purchased from Sigma (St. Louis, MO, USA). Polyethylene glycol (PEG) 6000 was purchased from Merck (Darmstadt, Germany) and NuncMaxisorb® 96 well immuno plates were obtained from Thermo Fisher Scientific (Roskilde, Denmark). AminoLink® matrix was purchased from Pierce (Rockford, IL, USA). 2, 2'-Azinobis [3-ethyl-2, 3-dihydrobenzithiazole-6-sulfonate] (ABTS) was obtained from Roche (Mannheim, Germany).

Reverse transcription PCR and quantitative PCR (RT-qPCR): TriFast® reagent for RNA extraction was obtained from PEQlab (Erlangen, Germany), transcriptor high fidelity cDNA synthesis kit from Roche (Mannheim, Germany), and DNAse 1 from Fermentas (Vilnius, Lithuania). The 2x SensiMix SYBR® Green Kit for qPCR was obtained from Quantace (Watford, England). The Rotor-Gene® 6000 real time thermal cycler with Rotor-gene® 6000 series software was from Corbett life science (Concorde, NSW, Australia).

3.2.2 METHODS

3.2.2.1 Recombinant expression and purification of full length *Tco*OPB

Terrific broth [12 g/l tryptone, 24 g/l yeast extract, 0.4% (v/v) glycerol, 2.31 g/l KH_2PO_4 and 12.54 g/l K_2HPO_4; 200 ml] (Tartof and Hobbs, 1987), containing 34 µg/ml of kanamycin was inoculated with a colony of recombinant *E. coli* BL21 DE3 cells containing the full length *T. congolense* oligopeptidase B gene (*TcoOPB*) in the pET28a expression vector and grown overnight at 37°C. Bacterial cells were harvested by centrifugation (5000 *g*, 15 min, RT) and resuspended in lysis buffer [PBS pH 7.4, Triton X-100 1% (v/v), 1 mg/ml lysozyme] at RT for 10 min before freezing at -20°C. Frozen samples were thawed and sonicated 4 times for 10 s each and separated by centrifugation (3000 *g*, 15 min, RT). The supernatant containing the soluble fraction and the pellet containing the insoluble fraction were separated on a 10% SDS-PAGE gel to confirm expression alongside untransformed *E. coli* BL21 DE3 cells for comparison. His-select® nickel affinity resin (1 ml) was placed in a 10 ml chromatography column, washed with 2 column volumes of deionised water and equilibrated with 5 column volumes of equilibration buffer (50 mM NaH_2PO_4, 500 mM NaCl, 10 mM imidazole, pH 6.8). Bacterial lysate (10 ml) containing soluble recombinant *Tco*OPB was added to the equilibrated resin and mixed using an end-over-end rotator at 4°C overnight. The unbound fraction was collected and the resin allowed to settle before washing with equilibration buffer until an A_{280} reading of < 0.02 was obtained. The bound fraction was recovered using 10 ml of elution buffer (50 mM NaH_2PO_4, 500 mM NaCl, 250 mM imidazole, pH 8). The nickel resin column was regenerated using 2 column volumes of deionised water, 5 column volumes of 6 M guanidine-HCl, another 3 column volumes of deionised water and a final 3-column volume wash with equilibration buffer before storage at 4°C in 30% (v/v) ethanol. The unbound fraction, washes and eluted fraction were analysed on a 10% SDS-PAGE gel for efficiency of purification and purity of recombinant *Tco*OPB. Purified recombinant *Tco*OPB fractions were pooled and concentrated using Amicon Centriprep® concentration filters according to the

manufacturers' specifications. The final concentration of *Tco*OPB was determined using dilutions of an albumin standard (doubling dilutions from 1.5 to 0.125 mg/ml) separated on a 10% SDS-PAGE gel alongside the purified protein sample.

3.2.2.3 Z-Arg-Arg-AMC hydrolysis by recombinant *T. congolense* oligopeptidase B

Serial dilutions of *Tco*OPB (0.5 to 0.0005 µg/ml) were made in 1 ml of 0.1% (w/v) Brij-35 and incubated with assay buffer [200 mM Tris-HCl buffer, pH 8, 10 mM DTT and 0.02% (w/v) NaN_3] for 10 min at 37°C. Aliquots were combined in triplicate with the Z-Arg-Arg-AMC substrate (20 µM) and the fluorescence measured (excitation at 360 nm and emission at 460 nm) using a FLUOstar OPTIMA fluorescence microplate reader (BMG Labtech, Offenburg, Germany). A sample comprising 0.1% (w/v) Brij-35 and assay buffer without enzyme was used as a negative control.

3.2.2.4 Anti-*Tco*OPB antibody preparation and ELISA

Anti-*Tco*OPB antibodies were raised in chickens. Briefly, four chickens were immunised intramuscularly at two different sites in the breast muscle at week 0 with 100 µg of recombinant *Tco*OPB emulsified with Freund's complete adjuvant in a 1:1 (v/v) ratio. Booster immunisations were performed at weeks 2, 4, 6 and 10 with 100 µg with *Tco*OPB emulsified in Freund's incomplete adjuvant in a 1:1 (v/v) ratio. Eggs were collected prior to the first immunisation to serve as pre-immune controls. Eggs were collected for a period of 14 weeks and stored at 4°C. Chicken immunoglobulin Y (IgY) was isolated from the eggs as previously described (Goldring and Coetzer, 2003). Briefly, the egg yolk was separated from the albumin, removed from the yolk sac and the yolk volume measured. Twice the yolk volume of isolation buffer [100 mM NaH_2PO_4, 0.02% (w/v) NaN_3, pH 7.6] was added and 3.5% (w/v) of PEG 6000 dissolved by gentle stirring. The solution was centrifuged (4420 g, 30 min, RT) and filtered through absorbent cotton wool. The PEG 6000 concentration was increased to 12% (w/v) by the addition of 8.5% (w/v) of PEG, dissolved and centrifuged as before to pellet the IgY. The pellet was re-dissolved in a volume of isolation buffer equal to the original yolk volume and 12% (w/v) PEG was added, dissolved and the final pellet collected by centrifugation (12 000 *g*, 10 min, RT). The final IgY pellet was dissolved in $1/6^{th}$ of the original yolk volume using isolation buffer containing 0.1% (w/v) NaN_3. The IgY concentration was determined by measuring the A_{280} of a 1:50 dilution of IgY and using an extinction coefficient of $E^{1\ mg/ml}_{280\ nm}=1.25$ (Goldring and Coetzer, 2003).

Recombinant *Tco*OPB was coupled to AminoLink® coupling gel for the affinity purification of antibodies produced against the protein. AminoLink® coupling gel was diluted 1:3 in coupling buffer [100 mM NaH_2PO_4, 0.05% (w/v) NaN_3, pH 7.4] to a total of 4 ml, pipetted into a chromatography column, washed three times with 10 ml of coupling buffer and allowed to settle. Recombinant *Tco*OPB (3 mg, purified as described in Section 3.2.2) was added to the column with 50 µl of 1 M $NaCNBH_3$, mixed on an end-over-end rotator (2 h, RT) and left to stand (4 h, RT). The gel was washed with one column volume of quenching buffer (1 M Tris-HCl, pH 7.4) and incubated on an end-over-end rotator (30 min, RT) with one column volume of quenching buffer containing 50 mM $NaCNBH_3$. The coupled gel was allowed to settle and washed with washing buffer (1 M NaCl) until an A_{280} reading of zero was obtained followed by a final wash with coupling buffer before storage at 4ºC. Isolated IgY was pooled into three fractions comprising of pool 1 (weeks 2 - 6 for all 4 immunised chickens), pool 2 (weeks 7 - 12 for chickens 1 and 3) and pool 3 (weeks 7 -12 for chickens 2 and 4) and filtered through Whatman no. 1 filter paper. The *Tco*OPB affinity column was equilibrated with washing buffer [100 mM NaH_2PO_4, 0.05% (w/v) NaN_3, pH 6.5] and each IgY pool cycled separately through the column overnight (10 ml/h, RT) to allow antibody binding to the immobilised recombinant protein. The column was washed with washing buffer until an A_{280} nm reading of ~ 0 was achieved to remove unbound IgY. Bound IgY was eluted with low pH elution buffer [100 mM glycine-HCl, 0.02% (w/v) NaN_3, pH 2.8] in fractions of 900 µl in 1.5 ml tubes containing 100 µl of neutralisation buffer [1 M NaH_2PO_4, 0.02% (w/v) NaN_3, pH 8.5]. Elution was monitored at A_{280} nm using 900 µl of elution buffer and 100 µl of neutralisation buffer as a blank. Fractions containing affinity purified IgY were pooled and 10% (w/v) NaN_3 added to a final concentration of 0.1% (w/v) before storage at 4ºC. The affinity column was regenerated using 12 column volumes of washing buffer.

Antibody production in chickens immunised with *Tco*OPB was monitored using enzyme linked immunoadsorbent assays (ELISAs) as has been previously described (Huson *et al.*, 2009). Recombinant *Tco*OPB in PBS (1 µg/ml) was used to coat the wells of a Nunc Maxisorb® microtitre plate (150 µl/well) and left overnight at 4°C. ELISA plates were patted dry on tissue paper to remove residual coating solution and 200 µl/well of blocking buffer [0.5% (w/v) BSA-PBS] was added and incubated (1 h, 37°C) to prevent non-specific binding of antibodies. Plates were washed three times with 0.1% (v/v) Tween-20-PBS and patted dry again. IgY isolated from eggs collected weekly from all 4 chickens was serially diluted in blocking buffer from 100 µg/ml to 0.01 µg/ml and 100 µl/well of each dilution incubated in triplicate for 2h at

37°C. Affinity purified IgY from all three pools was also serially diluted in blocking buffer from 100 µg/ml to 0.0001 µg/ml and 100 µl/well of each dilution incubated in triplicate for 2h at 37°C. All plates were washed three times with 0.1% (v/v) Tween-20-PBS and patted dry before adding rabbit anti-chicken IgY-HRPO labelled secondary antibody diluted at 1: 20 000 in blocking buffer and incubated (120 µl/well, 1 h, 37°C). The plates were given a final wash with 0.1% (v/v) Tween-20-PBS, patted dry and ABTS/H_2O_2 chromogen-substrate solution [0.05% (w/v) ABTS, 0.0015% (v/v) H_2O_2, 0.15 M citrate-phosphate buffer, pH 5] added and left to develop in the dark for 20 min before the absorbance was read at 405 nm using a FLUOstar OPTIMA microplate reader (BMG Labtech, Offenburg, Germany).

.

3.2.2.5 Cloning conserved gene regions of *Tco*OPB and *Tb*OPB into the p2T7^{Ti} RNAi plasmid vector and transformation of parasites

Cloning was carried out as described before for *Tco*CATL and *Tb*CATL in Section 2.2.2. Briefly, a pellet of parasites either from *T. congolense* (strain IL3000) or *T. b. brucei* (strain Lister 427) was resuspended in 1 ml of PBS (pH 7.2), centrifuged (1000 *g*, 10 min, RT), and the supernatant removed. The pellet was dissolved in 150 µl TELT [50 mM Tris-HCl buffer, pH 8, 62.5 mM EDTA, pH9, 2.5 M LiCl, 4% (v/v) Triton X-100] and incubated for 5 min at RT. Phenol-chloroform [1:1 (v/v); 150 µl] was added and the mixture placed on a rocker for 5 min at RT. The aqueous and non-aqueous phases were separated by centrifugation (13 000 *g*, 5 min, RT), and the upper phase added to 300 µl of absolute ethanol and centrifuged (13 000 *g*, 5 min RT) to precipitate genomic DNA. The supernatant was poured out and the DNA pellet washed with 1 ml of absolute ethanol before a last round of centrifugation (13 000 *g*, 5 min, RT). The ethanol was poured out and the pellet of genomic DNA incubated at 37ºC for 15 min to evaporate excess ethanol. The dry pellet of DNA was dissolved in 100 µl of TE buffer [100 mM Tris-HCl buffer, pH 7.5, 10 mM EDTA] with RNase (2 mg/ml) and incubated at 37ºC for 45 min.

DNA sequences coding for *OPB* gene regions that are conserved between *T. brucei* [EMBL accession no. AF078916, (Morty *et al.*, 1999b)] and *T. congolense* [Contig_0000238 GeneDB database [http://www.genedb.org/genedb/tcongolense/, accessed 10-10-2007, (Huson, 2006)] (see Appendix 1) were selected to design primers that amplified a 489 bp fragment in *T. congolense* and a 492 bp fragment in *T. brucei.* The primers were designed to introduce an XbaI restriction site at each end of the PCR product for subcloning into the p2T7^{Ti} RNAi plasmid. Primers for the 489 bp region in the *TcoOPB* gene were forward (5'-

CTGTCTAGA**ATG**TACGCCATTGCAGTGCGC-3') with an XbaI restriction site (underlined) and a start codon (bold), and reverse (5'-CACTCTAGA**TTA**GTCCGTCTTGAGCTCCCG-3') with an XbaI restriction site (underlined) and a stop codon (bold). Primers for the 492 bp region in the *TbOPB* gene were forward (5'-GATTCTAGA**ATG**ATATATGCTATTGCGCAC-3') with an XbaI restriction site (underlined) and a start codon (bold), and reverse (5'-CACTCTAGA**TTA**GTCTGTCTTGAGTTCACG-3') with an XbaI restriction site (underlined) and a stop codon (bold). Extracted genomic DNA was used as a template for PCR as described in Section 2.2.2 and the resulting PCR products were ligated into the pGEM-T® vector (T-vector) using a ratio of 3:1 with 1 x ligation buffer and 1 U of T4 DNA ligase at 4ºC overnight. The ligation mix was transformed into competent *E. coli* JM109 cells using the Transform Aid™ bacterial transformation kit according to the manufacturer's guide. Recombinant T-vector clones were selected as described in Section 2.2.2 and sequenced [Segoli sequencing unit, International Livestock Research Institute (ILRI), Nairobi, Kenya].

Recombinant OPB T-vector plasmids were restricted using XbaI and the resulting 489 bp (*Tco*OPB) and 492 bp (*Tb*OPB) products were subcloned into the XbaI linearised p2T7^{Ti} RNAi vector. Briefly, 50 µl of extracted plasmid DNA was restricted with XbaI to release the T-vector cloned products and separated on a 1% (w/v) agarose gel in 1 x TAE buffer at 80 V. The 320 bp and 300 bp products were cut out of the gel and purified using the E.Z.N.A® gel extraction kit according to the manufacturer's manual and eluted in 30 µl of elution buffer. The p2T7^{Ti} RNAi vector was also restricted with XbaI, cleaned and concentrated using the DNA clean and concentrator kit before treatment with 1 U of SAP for dephosphorylation to avoid the empty vector from re-ligating. The dephosphorylated vector and restricted fragments were used in separate ligation reactions at a ratio of 1:3 (vector: insert) with 1 x ligation buffer and 1 U of T4 DNA ligase at 4ºC overnight before transforming into *E. coli* JM109 cells and plating on 2x YT plates containing 100 µg/ml ampicillin. Positive clones were identified using colony PCR with insert primers and confirmed with restriction using XbaI. Recombinant p2T7^{Ti}/*Tb*OPB, p2T7^{Ti}/*Tco*OPB, and p2T7^{Ti}/*Tco*TUB used as a positive control (10 µg each) were linearised with NotI and transformed into procyclic *T. brucei* T7T 29 and *T. congolense* TRUM 183 29-13 cultures respectively as previously described in Section 2.2.3. Positive trypanosome clones were selected for genomic DNA isolation and PCR using primers specific to bleomycin [forward (5'-ATGGCCAAGTTGACCAGTGCC-3') and reverse (5'-TGCACGCAGTTGCCGGCCGGG-3')] to confirm integration of the p2T7^{Ti}/*Tb*OPB, p2T7^{Ti}/*Tco*OPB and p2T7^{Ti}/*Tco*TUB plasmids.

3.2.2.6 Western blot analysis of PCR positive clones

A pellet of 2 x 10^6 parasites from each positively identified culture was washed twice with PBS (pH 7.2) and resuspended in 50 µl of lysis buffer [10 mM EDTA, 1% (v/v) Triton X-100, 20 mM Tris-HCl buffer, pH 7.2, containing 10 µM L-trans-epoxysuccinyl-leucylamido(4-guanidino)butane (E64)], an equal volume of reducing gel loading buffer was added and samples boiled for 10 min. Lysates generated from ~ 5 x 10^5 parasites (25 µl) were separated per well on a 10% SDS-PAGE gel (Laemmli, 1970) at 18 mA per gel, blotted on nitrocellulose (Towbin *et al.*, 1979) and probed with chicken anti-*Tco*OPB antibodies (50 µg/ml dilution). Rabbit anti-chicken IgG-HRPO labelled secondary antibodies were used at 1:10000. Blots were developed using 0.06% (w/v) 4-chloro-1-naphthol, 0.0015% (v/v) H_2O_2 in Tris buffered saline (20 mM Tris-HCl, 200 mM NaCl, pH 7.4).

3.2.2.7 RNA analysis of PCR positive clones

RNA was isolated from transformed trypanosome cultures using the TriFast™ reagent based on a method described by Chomczynski and Sacchi (1987). Briefly, 1 x 10^7 parasites were centrifuged (1000 *g*, 10 min, RT) and the pellet resuspended in 1 ml of TriFast® reagent after discarding the supernatant. Parasites were dissolved using a pipette and kept at room temperature for 5 min. Chloroform (0.2 ml) was mixed in for 15 s and incubated (10 min, RT). This was followed by centrifugation (12 000 *g*, 5 min, RT) to separate the mixture into a lower phenol-chloroform phase and an upper aqueous phase that contained the RNA. The upper aqueous phase was transferred into a new 1.5 ml RNase free tube and RNA precipitated by the addition of 0.5 ml isopropanol. The samples were kept on ice for 15 min before centrifugation (12 000 *g*, 10 min, 4ºC). An RNA pellet was observed at the bottom of the tube after the supernatant was removed and this was washed twice with 75% (v/v) ethanol followed by centrifugation (12 000g, 10 min, 4ºC). The ethanol was poured off, the pellet air dried and the RNA pellet dissolved in 100 µl of formamide before storage at -20ºC.

Extracted RNA was quantified by measuring the absorbance at 260 nm and separated on a 1.4% agarose gel as previously described (Pelle and Murphy, 1993). All RNA samples were treated with RNase free DNAse according to the manufacturers' specifications and used for generating complementary DNA (cDNA), the first step in the RT-PCR reaction, by using the high fidelity cDNA synthesis kit. Briefly, 1 µg of RNA as determined by spectrophotometry was combined with 0.5 mM dNTPs, 0.25 µM oligo $(dT)_{18}$, and DEPC (diethylpyrocarbonate)

treated water (0.1%, v/v) to a final volume of 10 µl and incubated at 65ºC for 10 min in a thermocycler. Samples were chilled on ice for 2 min before adding 4 µl of 5x reverse transcriptase buffer, 1 U of RNase inhibitor, 1 U of reverse transcriptase and DEPC treated water to a final volume of 20 µl and incubated at 40ºC for 50 min in a thermocycler before a final extension temperature at 70ºC for 15 min. cDNA generated was then used as a template for the second step of RT-PCR with gene specific primers for *TbOPB* and *T. congolense* tubulin (*TcoTUB*) (Table 3.1). A master mix for the reaction containing sets of primers designed for *Tb*OPB or *Tco*TUB (0.25 µM each) and *T. brucei* and *T. congolense* cDNA respectively, 1 x PCR buffer, 2.5 mM $MgCl_2$, 1 U Taq and 0.5 mM dNTPs in a total reaction volume of 25 µl was prepared. PCR amplification was started with an Taq polymerase activation step at 95ºCfor 5 min, followed by 25 cycles of a DNA denaturing step at 95ºC for 30 s, a primer annealing step at 55ºC for 30 s, and an elongation step at 72ºC for 1 min. A final elongation step was added at the end of 25 cycles at 72ºC for 7 min. The resulting PCR products were separated on a 1% (w/v) agarose gel.

Table 3.1 Primer sequences designed for quantitative PCR. Primers were used to analyse down-regulation of *Tb*OPB in *T. brucei* T7T parasites transformed with p2T7^{Ti}/*Tb*OPB. *T. congolense* TRUM 183 29-13 parasites, transformed with p2T7^{Ti}/*Tco*TUB, were used as a positive control for the experiment and actin as the reference gene for PCR.

Primer name	Sequence (5'→3')
*Tb*OPB forward	5'-ATGATATATGCTATTGCGCAC-3'
*Tb*OPB reverse	5'-GTCTGTCTTGAGTTCACG-3'
*Tb*ACT forward	5'-GACGAGGAACAAACTGCT-3'
*Tb*ACT reverse	5'-ACCGTCACCAGCGTCGAG-3'
*Tco*TUB forward	5'-ATGCGCGAGGCTATCTGCATC-3'
*Tco*TUB reverse	5'-GTAACTCCTCAACGTCCTCCTC-3'

T. brucei cDNA generated as described above was also used for quantitative PCR (qPCR). Briefly, primers for either *TbOPB* or *T. brucei* actin (*TbACT*) (0.25 µM each) were added to a master mix containing *T. brucei* cDNA template, 25 µl of 2x SensiMix dT [2x PCR buffer, 6 mM $MgCl_2$, 1 mM dNTPs and 2 U Taq], 1 µl of 50x SYBR® green solution and distilled water to a final volume of 50 µl. *T. brucei* cDNA generated from different concentrations of RNA ranging from 0 – 100 ng was used to measure the comparable cycling efficiency of *Tb*ACT and *Tb*OPB primers to confirm using them together in one experiment as described by Livak

and Schmittgen (2001). Relative gene expression was calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). This method operates by comparing the cycle threshold values of different PCR reactions. Relative concentrations of cDNA present during the exponential phase of the reaction are determined by plotting fluorescence against cycle number on a logarithmic scale. A threshold for detection of fluorescence above background is determined and the cycle at which the fluorescence from a sample crosses the threshold is called the cycle threshold, C_T. ΔC_T was calculated by subtracting the *Tb*ACT primer C_T values from *Tb*OPB primer C_T values obtained when using the same concentration of cDNA on different days of the experiment. $\Delta\Delta C_T$ is obtained by deducting the average C_T value of the whole experiment from each individual value to calculate the percent change in cDNA for each reaction. These values are then plotted as a percentage change over the time period of the experiment. Relative amounts of RNA for each clone in the subsequent days of the experiment were plotted in comparison to day one, which was taken at 100%. For statistical analyses, values were expressed as means ± standard error of the mean (SEM). Significance levels were calculated by unpaired t tests and differences were considered significant at a *p* value <0.001.

3.3 RESULTS

3.3.1 Recombinant expression and purification of *Tco*OPB

Expression of the pET28a-*Tco*OPB clone in *E. coli* BL21 DE3 cells was successful as evidenced by the presence of a band at ~80 kDa in both the insoluble (Fig. 3.1 A, lanes 2-4) and soluble (Fig. 3.1 A, lanes 5-6) fractions of the bacterial lysate. The soluble fraction was used for His-tagged affinity purification where most of the recombinant protein was successfully separated from bacterial lysate and eluted with 250 mM imidazole (Fig.3.1 B, lane 5). His-tagged purified *Tco*OPB fractions were pooled (Fig. 3.1 C, lanes 1 & 2) and concentrated (Fig. 3.1 C, lane 3) before quantification by comparison with BSA standards (Fig. 3.1 D). The concentration of *Tco*OPB was estimated to be between 0.75 and 1 mg/ml. Dilutions of purified *Tco*OPB was used for an activity assay with Z-Arg-Arg-AMC, its preferred substrate, to confirm that the purified enzyme was active (Fig. 3.2).

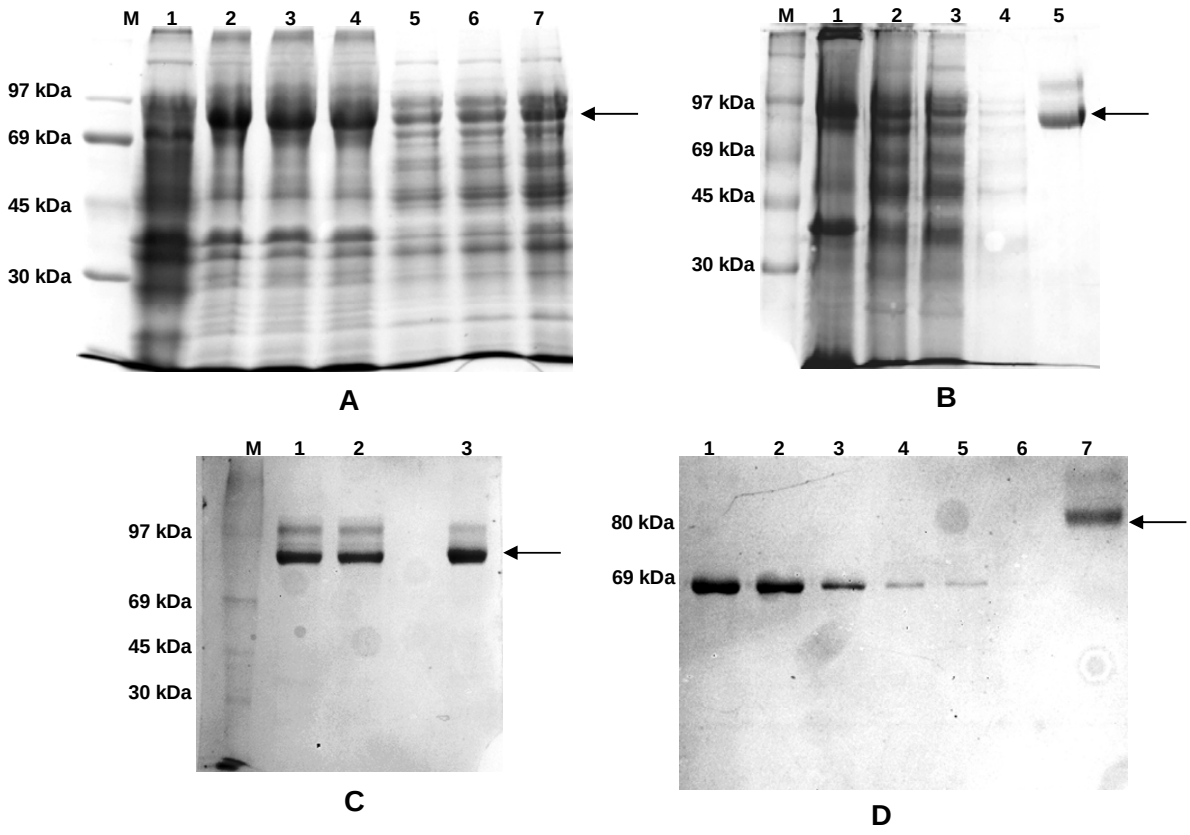


Figure 3.1 SDS-PAGE analysis of recombinant expression and affinity purification of full length *T. congolense* OPB. Samples were analysed on 10% SDS-PAGE gels. **M**: molecular weight markers. **A: lane 1**, untransformed *E. coli* BL21 DE3 lysate; **lanes 2-4**, insoluble fraction of *T. congolense* OPB expression; **lanes 5-7**, soluble fraction of *T. congolense* OPB expression; expected size of recombinant *T. congolense* OPB at ~ 80 kDa (←). **B: lane 1**, soluble fraction of *T. congolense* OPB expression loaded onto the HIS purification column; **lane 2**, unbound fraction; **lane 3**, first column wash; **lane 4**, second column wash; **lane 5**, eluted *T. congolense* OPB fraction. **C: lanes 1 & 2**, dilute *T. congolense* OPB fractions; **lane 3**, concentrated *T. congolense* OPB. **D:** BSA quantification of purified recombinant *T. congolense* OPB. By loading 10 µl/well of the following BSA dilutions; **lane 1**, 1.5 mg/ml; **lane 2**, 1 mg/ml; **lane 3**, 0.75 mg/ml; **lane 4**, 0.5 mg/ml; **lane 5**, 0.25 mg/ml; **lane 6**, 0.125 mg/ml; **lane 7**, purified and concentrated *T. congolense* OPB fraction estimated to be between 1 – 0.75 mg/ml.

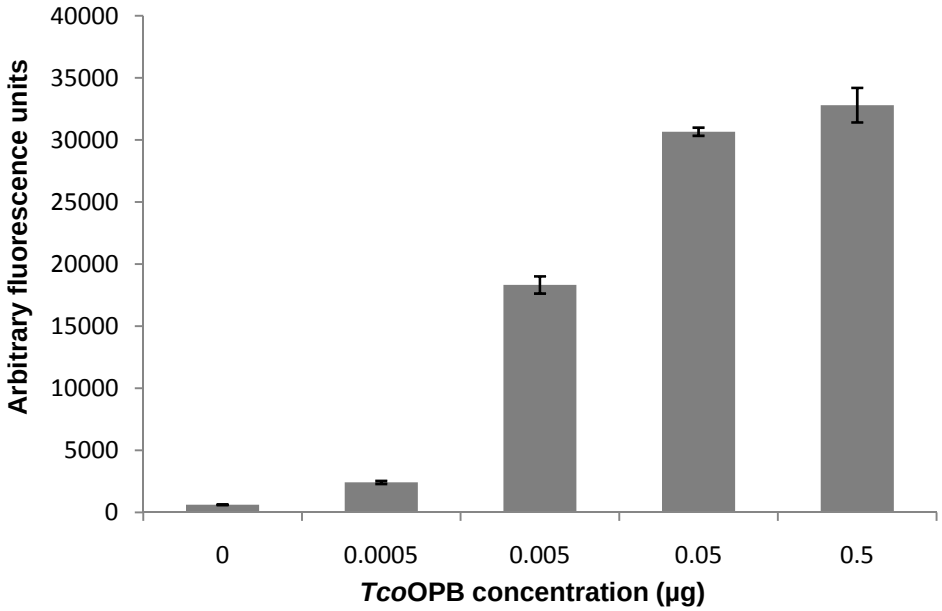


Figure 3.2 Purified recombinant oligopeptidase B from *T. congolense* hydrolyses Z-Arg-Arg-AMC. Different concentrations of recombinant *Tco*OPB were used in an activity assay with 0.1% (w/v) Brij-35. A mean of triplicate fluorescence values was plotted as arbitrary fluorescence units ± SEM.

3.3.2 Production of anti-*Tco*OPB IgY antibodies

Purified *Tco*OPB was used to raise antibodies in four chickens and the antibodies were isolated from the egg yolks. IgY antibodies from all four chickens were analysed over a 14-week period using ELISA to determine when antibody production peaked (Fig. 3.3 A). Chicken 1 produced the highest titre antibodies in weeks 6 and 7 after the first immunisation (Fig. 3.3 B, *Tco*OPB 1), while chicken 2 produced antibodies with a lower titre that peaked at week 5 (Fig. 3.3 B, *Tco*OPB 2). The antibodies produced by chicken 3 also peaked at week 5 (Fig. 3.3 B, *Tco*OPB 3) and those produced by chicken 4 at week 6 (Fig. 3.3 B, *Tco*OPB 4).

IgY was pooled into three different groups comprising of pool 1 (weeks 2 - 6 for all immunised chickens), pool 2 (weeks 7 - 12 for chickens 1 and 3) and pool 3 (weeks 7 -12 for chickens 2 and 4) and affinity purified using the recombinant *Tco*OPB affinity column (Fig. 3.4, A). ELISAs using recombinant OPB and affinity purified antibodies showed that pool 3 had the highest titre of the three pools (Fig. 3.4, B).

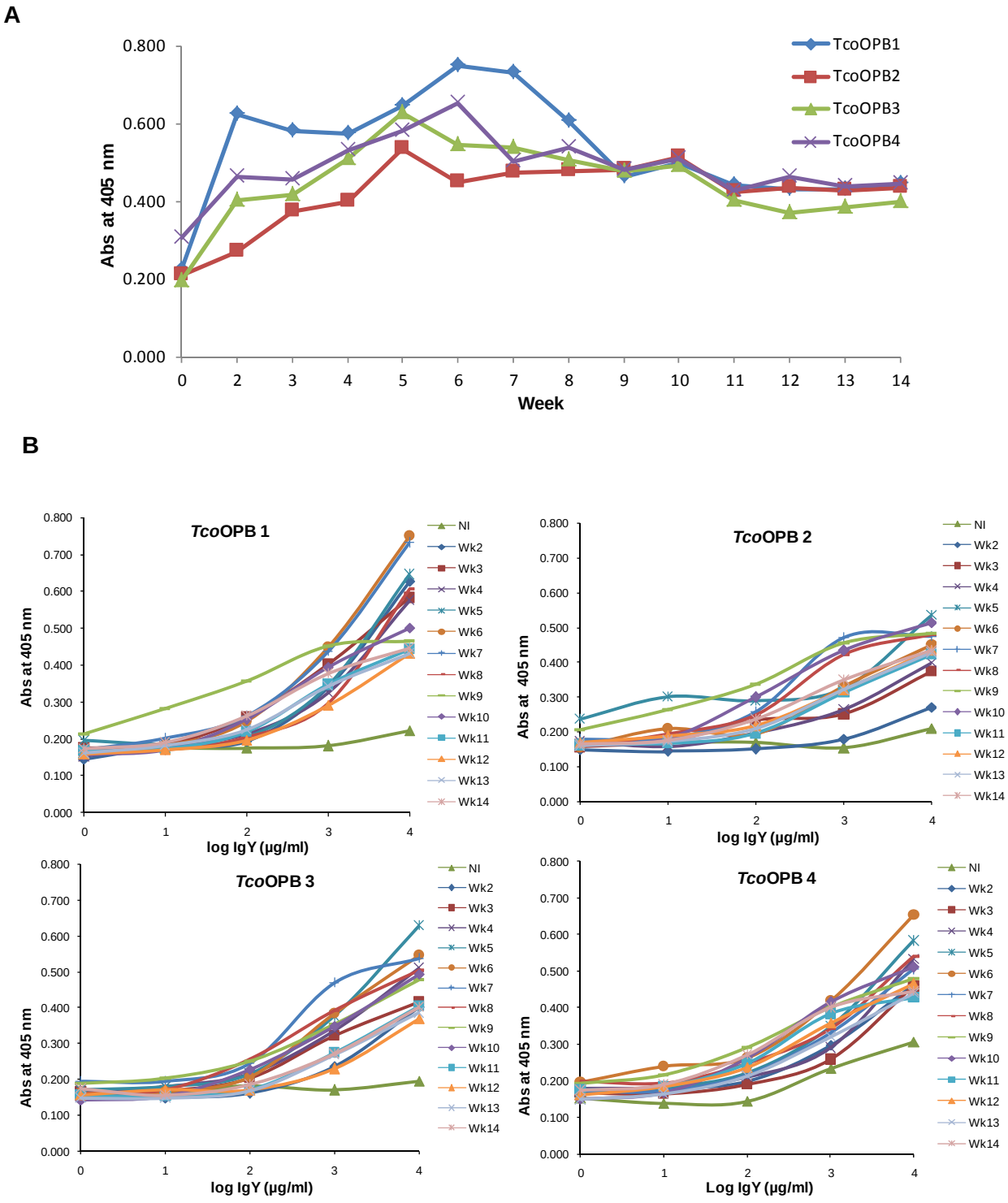


Figure 3.3 ELISAs showing antibody production in chickens following immunisation with recombinant *T. congolense* OPB. A: ELISA of IgY from chickens 1-4 showing antibody production over a 14-week period. **B: *Tco*OPB 1,** titration of antibodies produced by chicken 1; ***Tco*OPB 2,** titration of antibodies produced by chicken 2; ***Tco*OPB 3,** titration of antibodies produced by chicken 3; ***Tco*OPB 4,** titration of antibodies produced by chicken 4. ELISA plates were coated with 150 µl (1 µg/ml) of recombinant *Tco*OPB per well. NI, non-immune control IgY.

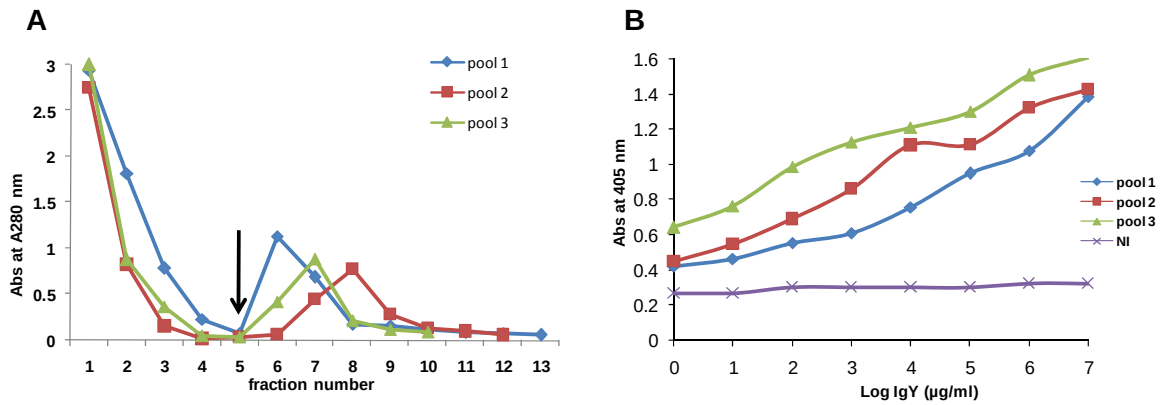


Figure 3.4 Affinity purification of anti-*Tco*OPB IgY and evaluation by ELISA. A: Elution profiles for affinity purification of anti-*Tco*OPB IgY. **Pool 1**, weeks 2 - 6 for all immunised chickens; **pool 2**, weeks 7 - 12 for chickens 1 and 3; **pool 3**, weeks 7 -12 for chickens 2 and 4. Arrow indicates point at which elution with low pH glycine-HCl buffer was started. **B:** ELISA evaluation of anti-*Tco*OPB affinity purified antibody pools 1-3. **NI**, non-immune control IgY. The ELISA plate was coated with 150 µl (1 µg/ml) of recombinant *Tco*OPB per well. Each point plotted is the average of 3 replicates.

Affinity purified chicken anti-*Tco*OPB IgY pools were characterised using a western blot with recombinant *Tco*OPB and lysates generated from *T. congolense* IL3000 and *T. brucei* 427 Lister procyclic cultures (Fig. 3.5). Affinity purified pool 3 anti-*Tco*OPB antibodies recognised affinity purified recombinant *Tco*OPB (Fig. 3.5 A, lane 1) but did not recognise native OPB in lysates generated from *T. congolense* (Fig. 3.5 B, lane 1) or *T. brucei* (Fig. 3.5 B, lane 2).

As an alternative to using IgY raised from full length recombinant *Tco*OPB, anti-*Tco*OPB peptide antibodies (chicken anti-*Tco*OPB peptide antibody, weeks 4 -7, 50 µg/ml) raised in a previous study (Huson, 2006) were used in a western blot. These antibodies recognised affinity purified recombinant *Tco*OPB (Fig. 3.5 C, lane 1) and OPB in both *T. congolense* and *T. brucei* lysates (Fig. 3.5 D, lanes 1 and 2 respectively).

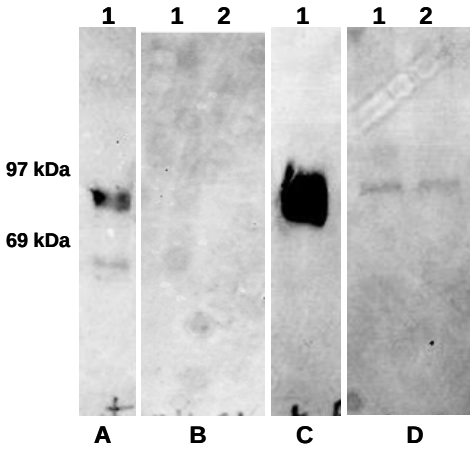


Figure 3.5 Western blot analysis of recombinant *Tco*OPB and parasite lysates with affinity purified chicken anti-*Tco*OPB IgY antibodies. A: lane 1, recombinant *Tco*OPB probed with chicken anti-*Tco*OPB IgY (50 µg/ml); **B: lane 1**, *T. congolense* parasite lysate and **lane 2**, *T. brucei* parasite lysate, probed with chicken anti-*Tco*OPB IgY (50 µg/ml). **C: lane 1**, recombinant *Tco*OPB probed with anti-*Tco*OPB peptide IgY (50 µg/ml). **D: lane 1**, *T. congolense* parasite lysate and **lane 2**, *T. brucei* parasite lysate, probed with anti-*Tco*OPB peptide IgY (50 µg/ml).

3.3.3 Cloning conserved gene regions of *Tco*OPB and *Tb*OPB into the RNAi vector p2T7^{Ti}

Genomic DNA was isolated from *T. brucei* 427 Lister and *T. congolense* IL300 strains (Fig. 3.6 A1, lanes 1 and 2 respectively) and used as a template for genomic DNA PCR with primers that had been designed to amplify conserved regions of the *OPB* gene (Fig.3.6 A2). A 489 bp product from *T. congolense* IL3000 (*TcoOPB*; lane 1) and a 492 bp product from *T. brucei* 427 Lister (*TbOPB*; lane 2) were amplified and ligated with the pGEM® T-vector. Blue-white screening and colony PCR revealed six T-vector recombinant clones for *TcoOPB* of 489 bp (Fig. 3.6 B1, lanes 1-6) and ten for *TbOPB* of 492 bp (Fig.3.6 B1, lanes 1-10). Plasmid mini-preparations for *TbOPB* (Fig. 3.6 C, lane 1) and *TcoOPB* (Fig. 3.6 C, lane 2) in T-vector were prepared. The T-vector constructs for *TcoOPB* and *TbOPB* were confirmed using XbaI restriction releasing a 489 bp fragment for *TcoOPB* and 492 bp fragments from two clones of *TbOPB* (Fig. 3.6 D1, lane 1 and Fig. 3.6 D2, lanes 1 and 2 respectively).

The XbaI linearised p2T7^{Ti} RNAi vector of 5 kb (Fig. 3.7 A, lane 1) was ligated with the 489 bp *TcoOPB* and the 492 bp *TbOPB* (Fig. 3.7 A, lanes 2 and 3 respectively). Recombinant clones were selected for colony PCR using insert primers that amplified 489 bp products for recombinant p2T7^{Ti}/*Tco*OPB (Fig. 3.7 B1, lanes 2, 3 and 9) and 492 bp products for recombinant p2T7^{Ti}/*Tb*OPB (Fig. 3.7 C1, lanes 1, 2, 6 and 9-11).

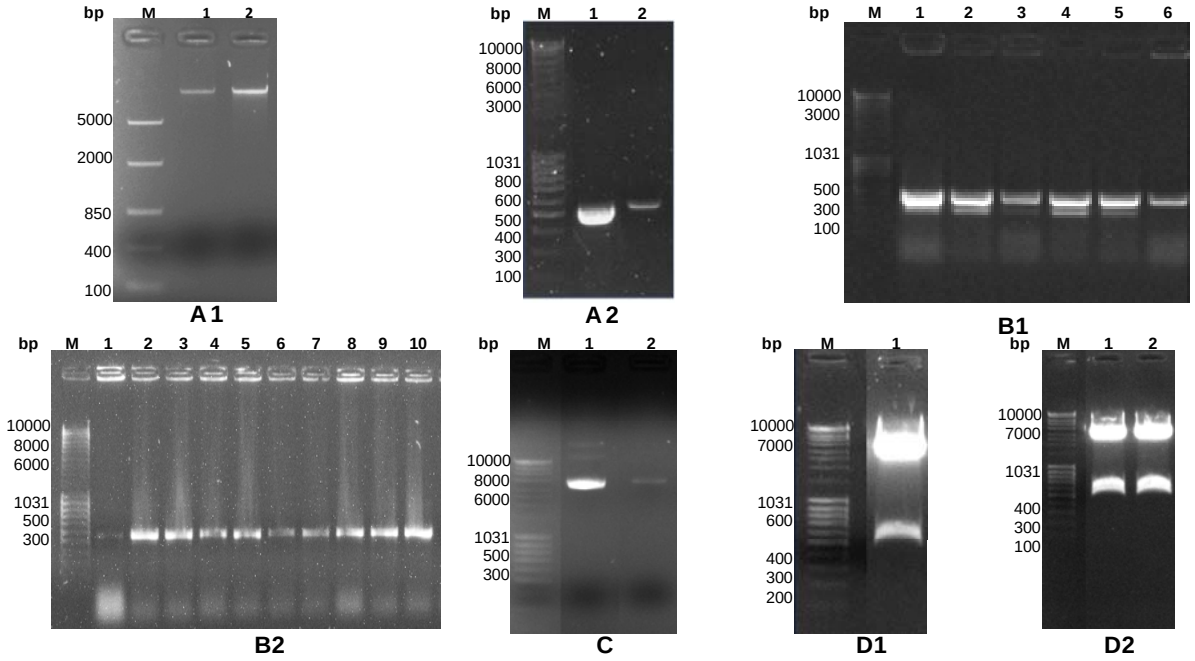


Figure 3.6 Amplification and T-vector cloning of conserved OPB gene regions from *T. brucei* and *T. congolense*. M, MassRuler® DNA ladder mix / Middle Range™ DNA ladder mix. **A1:** Genomic DNA isolation; **lane 1**, *T. brucei* 427 Lister; **lane 2**, *T. congolense* IL3000. **A2: lane 1**, 489 bp *TcoOPB* PCR product using primers designed for *T. congolense* genomic DNA; **lane 2**, 492 bp *Tb*OPB PCR product using primers designed for *T. brucei* genomic DNA. **B1:** Colony PCR **lanes 1-6**, *Tco*OPB clones. **B2:** Colony PCR **lanes 1-10**, *Tb*OPB clones. **C:** T-vector miniprep **lane 1**, *TbOPB*; **lane 2**, *Tco*OPB. **D1: lane 1**, XbaI restriction of *Tco*OPB in the T-vector. **D2: lanes 1-2**, XbaI restriction of *Tb*OPB in the T-vector. All samples were separated on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml)**.**

Successful subcloning into the p2T7^{Ti} RNAi vector was confirmed using XbaI restriction for p2T7^{Ti}/*Tco*OPB, releasing 489 bp products from three clones (Fig. 3.7 B2, lanes 1-3) and p2T7^{Ti}/*Tco*OPB, releasing 492 bp products from six clones (Fig. 3.7 C2, lanes 1-6). Recombinant p2T7^{Ti}/*Tco*OPB and p2T7^{Ti}/*Tb*OPB clones were prepared for transformation into *T. congolense* TREU 29-13 and *T. brucei* T7T-29 procyclic parasites by restriction with NotI giving *ca.* 5.4 kb fragments (Fig. 3.7 D, lanes 1 and 2 respectively). DNA isolated from a p2T7^{Ti}/*Tco*TUB clone was also restricted with NotI as a control for transformation (Fig. 3.7 D, lane 3).

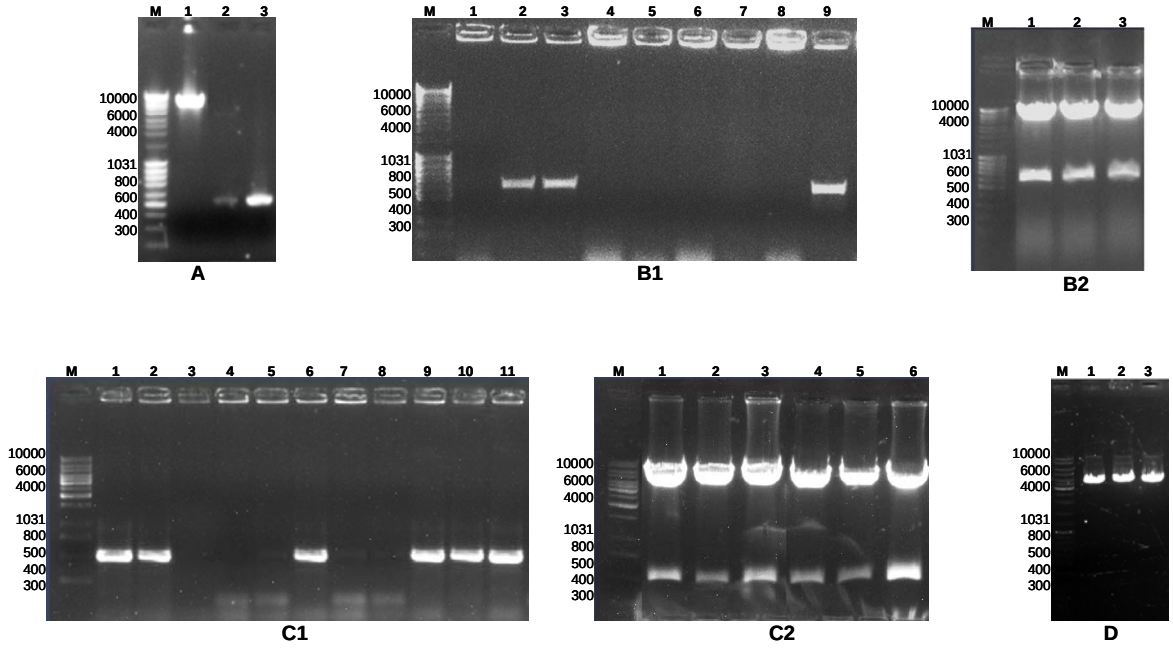


Figure 3.7 Subcloning conserved gene regions from *T. brucei* OPB and *T. congolense* OPB into the p2T7^{Ti} RNAi vector. M, MassRuler® DNA ladder mix in all gels. **A: lane 1**, gel purified XbaI restricted p2T7^{Ti} vector; **lane 2**, XbaI restricted 489 bp gel purified *Tco*OPB fragment from T-vector; **lane 3**, XbaI restricted 492 bp gel purified *TbOPB* fragment from T-vector. **B1: lanes 2, 3 and 9**, positive p2T7^{Ti}/*Tco*OPB clones using 489 bp primers; **B2: lanes 1-3**, p2T7^{Ti}/OPB clones restricted with XbaI. **C1: lanes 1, 2, 6 and 9-11**, positive p2T7^{Ti}/*Tb*OPB clones using 492 bp primers. **C2: lanes 1-6**, p2T7^{Ti}/*Tb*OPB clones restricted with XbaI. **D: lane 1**, p2T7^{Ti}/*Tco*OPB clone restricted with NotI; **lane 2,** p2T7^{Ti}/*Tb*OPB clone restricted with NotI; **lane 3**, p2T7^{Ti}/*Tco*TUB clone restricted with NotI in preparation for transformation. Products were separated on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml)**.**

3.3.4 Generation and characterisation of *T. congolense* and *T. brucei* p2T7^{Ti}/OPB clones

T. congolense p2T7^{Ti}/*Tco*OPB, *T. congolense* p2T7^{Ti}/*Tco*TUB and *T. brucei* p2T7^{Ti}/*Tb*OPB positive clones were expanded, used for genomic DNA isolation and analysed in a PCR reaction with bleomycin primers. The two sets of bleomycin primers used resulted in amplification of 350 bp (Fig. 3.8A, lanes 1-6) and 1000 bp (Fig. 3.8 A, lanes 7-12) products. Anti-*Tco*OPB peptide IgY was used in western blots to detect OPB in lysates from *T. congolense* p2T7^{Ti}/*Tco*OPB and *T. brucei* p2T7^{Ti}/*Tb*OPB positive clones. A three-day RNAi induction experiment using 1 µg/ml of tetracycline in *T. congolense* p2T7^{Ti}/*Tco*OPB cultures (Fig. 3.8 B, lanes 4-6) did not show any difference in OPB expression when compared to non-induced cultures used as a control (Fig. 3.8 B, lanes 1-3).

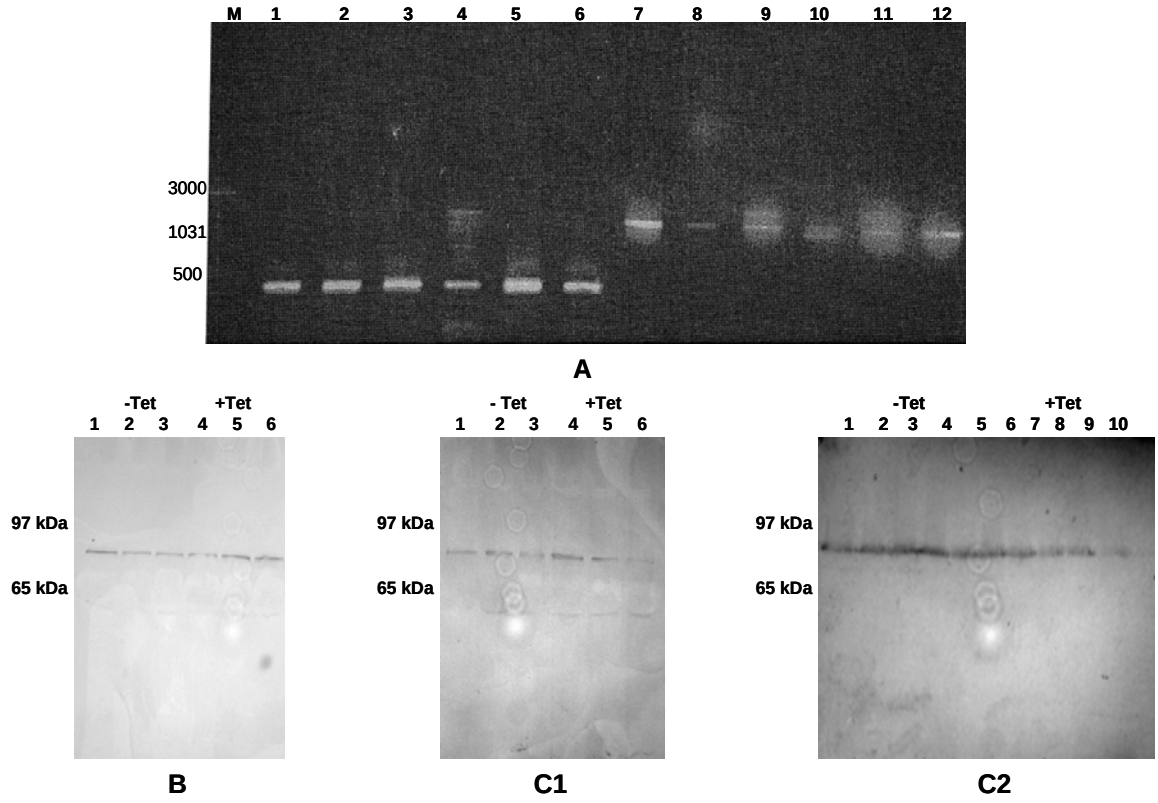


Figure 3.8 Evaluation of clones transformed for RNAi using PCR and western blotting. A: Genomic DNA PCR on *T. brucei* T7T and T. congolense TRUM 183 clones transformed with p2T7^{Ti}/OPB and p2T7^{Ti}/TUB. M, MassRuler® DNA ladder mix; 350 bp bleomycin primers with: **lane 1,** p2T7^{Ti}/BIP plasmids as a PCR control; **lane 2**, *T. congolense* p2T7^{Ti}/*TUB* genomic DNA; **lane 3**, *T. brucei* p2T7^{Ti}/OPB clone 1 genomic DNA; **lane 4**, *T. brucei* p2T7^{Ti}/OPB clone 2 genomic DNA; **lane 5**, *T. congolense* p2T7^{Ti}/OPB clone 1 genomic DNA; **lane 6,** *T. congolense* p2T7^{Ti}/OPB clone 2 genomic DNA. 1000 bp bleomycin primers with: **Lane 7,** p2T7^{Ti}/BIP plasmid as a PCR control; **lane 8**, *T. congolense* p2T7^{Ti}/*TUB* genomic DNA; **lane 9**, *T. brucei* p2T7^{Ti}/OPB clone 1 genomic DNA; **lane 10**, *T. brucei* p2T7^{Ti}/OPB clone 2 genomic DNA; **lane 11**, *T. congolense* p2T7^{Ti}/OPB clone 1 genomic DNA; **lane 12**, *T. congolense* p2T7^{Ti}/OPB clone 2 genomic DNA. Products were separated on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml). **B: Western blot analysis of three-day RNAi induction of *T. congolense* p2T7^{Ti}/OPB transformed clones using anti-*Tco*OPB peptide IgY. Lanes 1-3**, days 1-3 without induction (-Tet); **lanes 4-6**, days 1-3 with induction (+Tet). **C1: Western blot analysis of three-day RNAi induction of *T. brucei* p2T7^{Ti}/OPB transformed clones using anti-*Tco*OPB peptide IgY. Lanes 1-3**, days 1-3 without induction (-Tet); **lanes 4-6**, days 1-3 with induction (+Tet). **C2: Western blot analysis of 14-day RNAi induction of *T. brucei* p2T7^{Ti}/OPB transformed clones using anti-*Tco*OPB peptide IgY. Lanes 1-5**, days 1, 3, 6, 9 and 14 respectively without induction (-Tet); **lanes 6-10**, days 1, 3, 6, 9 and 14 respectively with induction (+Tet).

A parallel three-day induction experiment with tetracycline to induce RNAi in p2T7^{Ti}/*Tb*OPB clones revealed the down-regulation of OPB (Fig. 3.8 C1, lanes 4 - 6) when compared to the non-induced culture (Fig. 3.8 C1, lanes 1 - 3). OPB expression was, however, still noticeable after 3 days and a longer 14-day tetracycline induction experiment was carried out (Fig. 3.8

C2, lanes 6 - 10). OPB expression levels where still detected after 14 days, but at lower levels than in the non-induced cultures (Fig. 3.8 C2, lanes 1 - 5).

In order to observe how efficiently RNAi was functioning, RNA was isolated for RT-PCR analysis of p2T7^{Ti}/*Tb*OPB clones grown for three days either in the absence of tetracycline (-Tet) (Fig. 3.9 A, lanes 7 - 9) or in the presence of tetracycline (+Tet) (Fig. 3.9 A, lanes 10 - 12). The p2T7^{Ti}/*Tco*TUB culture was also used for RNA isolation to serve as a positive control for RNAi and grown for three days without tetracycline (-Tet) (Fig. 3.9 A, lanes 1 - 3) or induced with tetracycline (+Tet) (Fig. 3.9 A, lanes 4 - 6).

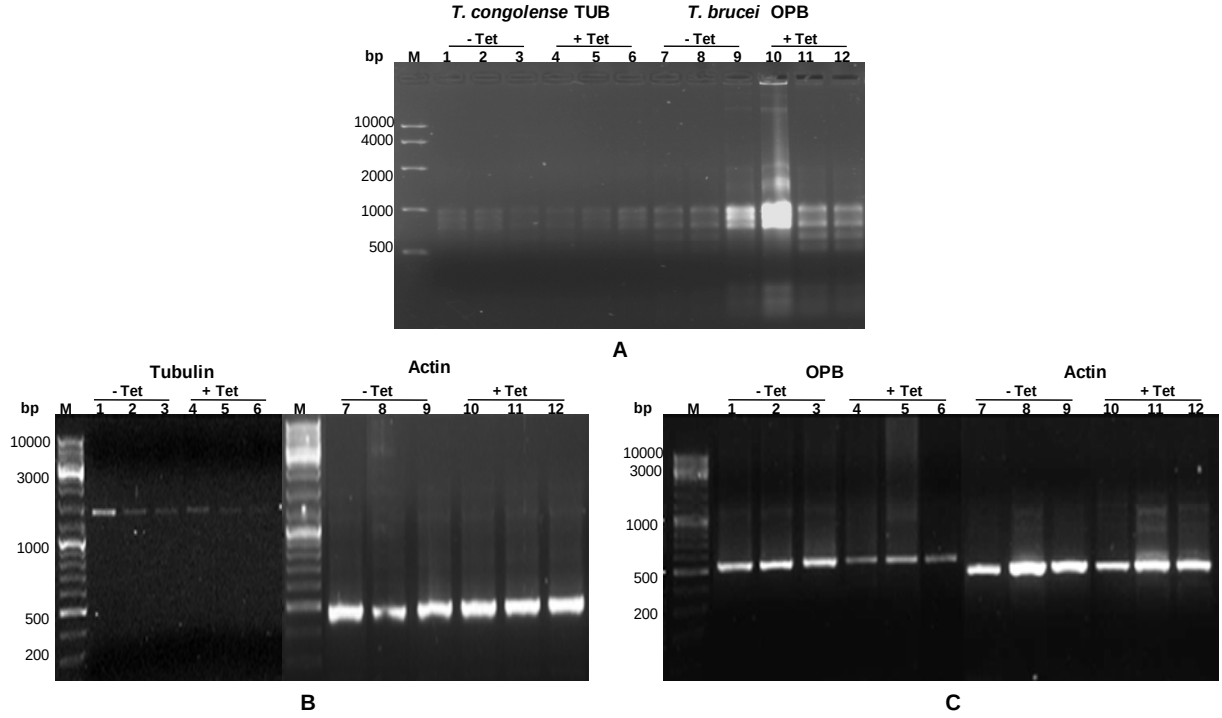


Figure 3.9 Three-day RNAi induction RT-PCR analysis of *T. congolense* p2T7^{Ti}/TUB and *T. brucei* p2T7^{Ti}/OPB transformants. A: RNA isolation. *T. congolense* p2T7^{Ti}/TUB days 1-3: **lanes 1-3**, with no induction (-Tet); **lanes 4-6**, with induction (+Tet). *T. brucei* p2T7^{Ti}/OPB days 1-3: **lanes 7-9**, with no induction (-Tet); **lanes 10-12**, with induction (+Tet). **B: RT-PCR analysis of *T. congolense* p2T7^{Ti}/TUB transformants.** Days 1-3: **lanes 1-3**, with no induction (-Tet) and **lanes 4-6**, with induction (+Tet) using 1200 bp tubulin primers; **lanes 7-9**, with no induction (-Tet) and **lanes 10-12**, with induction (+Tet) using 400 bp actin primers. **C: RT-PCR analysis of *T. brucei* p2T7^{Ti}/OPB transformants,** days 1-3: **lanes 1-3**, with no induction (-Tet) and **lanes 4-6**, *T. brucei* p2T7^{Ti}/OPB days 1-3 with induction (+Tet) using 492 bp *Tb*OPB primers; **lanes 7-9**, with no induction (-Tet) and **lanes 10-12**, with induction (+Tet) using 400 bp actin primers.

RT-PCR analysis revealed that RNAi induction in p2T7^{Ti}/*Tco*TUB control clones over a three-day period was efficient with a clear reduction in levels of tubulin in cultures induced with

tetracycline (Fig. 3.9 B, lanes 4 - 6) when compared to non-induced cultures (Fig. 3.9 B, lanes 1 - 3). Actin primers were used as a loading control for both RNAi induced and non-induced cultures (Fig. 3.9 B, lanes 7 - 12). Levels of OPB in *T. brucei* cultures when RNAi was induced with tetracycline for three days were only slightly reduced (Fig. 3.9 C, lanes 4 - 6) when compared to non-induced cultures (Fig. 3.9 C, lanes 1 - 3). Actin primers were also used as loading controls for both RNAi induced and non-induced cultures (Fig. 3.9 C, lanes 7 - 12). A repeat of the experiment over a 14-day period using RNA from tetracycline induced (Fig. 3.10, A, lanes 6 -10) and non-induced (Fig. 3.10, A, lanes 1 - 5) p2T7^{Ti}/*Tb*OPB cultures revealed a steady but incomplete reduction of OPB RNA in tetracycline induced clones (Fig. 3.10, C, lanes 1 – 5). A reduction was also observed in non-induced cultures over the same 14-day period (Fig. 3.10, B, 1 - 5). Actin primers were used as loading controls for both induced and non-induced p2T7^{Ti}/*Tb*OPB cultures (Fig. 3.10, C, lanes 6 -10 and Fig. 3.10, B, lanes 6 -10). The actin loading control in lane 9 of Fig. 3.1 B suggests that less sample was loaded affecting the amount of product seen in lane 4 of Fig. 3.1 B.

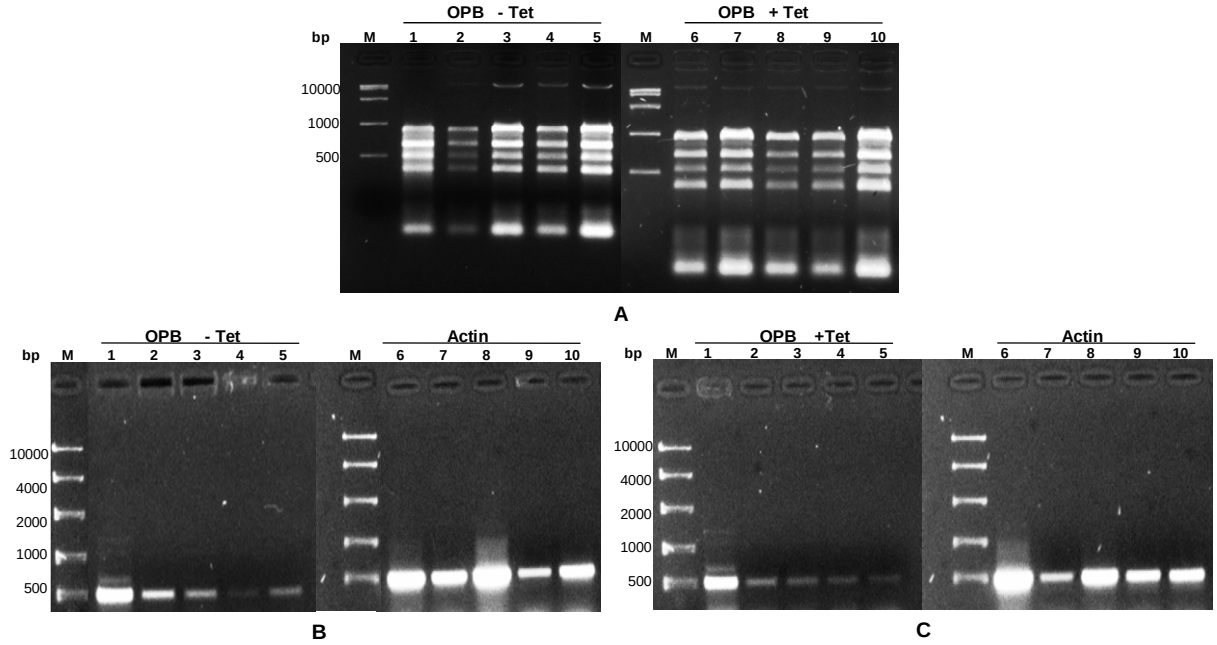


Figure 3.10 RT-PCR analysis of 14-day induction RNAi induction in *T. brucei* p2T7^{Ti}/OPB transformants. A: RNA isolation from *T. brucei* p2T7^{Ti}/OPB. Lanes 1-5 no induction (-Tet) days 1, 3, 6, 9 and 14 respectively; **lanes 6-10**, induction (+Tet) days 1, 3, 6, 9 and 14. **B: RT-PCR analysis of *T. brucei* p2T7^{Ti}/OPB transformants with no induction (-Tet). Lanes 1-5**, days 1, 3, 6, 9 and 14 using 492 bp *Tb*OPB primers; **lanes 6-10**, days 1, 3, 6, 9 and 14 using 400 bp actin primers. **C: RT-PCR analysis of *T. brucei* p2T7^{Ti}/OPB transformants with induction (+Tet). Lanes 1-5**, days 1, 3, 6, 9 and 14 using 492 bp oligopeptidase B primers; **lanes 6-10**, days 1, 3, 6, 9 and 14 using 400 bp actin primers.

In order to quantify the relative amounts of *Tb*OPB that were targeted for RNAi in selected clones, RNA was isolated for from clones grown for 14 days either in the absence (-Tet) (Fig. 3.10 A, lanes 1-5) or the presence (+Tet) (Fig. 3.10 A, lanes 6-10) of tetracycline. An initial experiment that compared the thermo-cycling efficiencies of OPB and actin primers at different RNA concentrations was performed as previously described (Livak and Schmittgen, 2001) (Fig. 3.11 A).

A plot of cDNA dilutions versus cycling efficiencies (ΔC_T) produced a line with a slope of 0.0397 meaning that *Tb*OPB primers were operating at a similar efficiency to *T. brucei* actin primers (*Tb*ACT). RNA at a concentration of 10 ng was used to calculate the relative reduction of *Tb*OPB levels in tetracycline induced cultures compared to non-induced cultures over a 10-day period (Fig. 3.11 B). The relative levels of RNA from tetracycline-induced cultures were reduced to *ca.* 50% by day five when compared to day one and this trend in reduction was observed until day 14. *Tb*OPB RNA levels at day nine of RNAi induction was calculated at *ca.*13% when compared to day one (Fig. 3.11 B). A reduction in RNA was also observed in non-induced cultures with day nine giving the lowest level at *ca.* 40% of RNA when compared to day one (Fig. 3.11 B).

The number of parasites in *T. congolense* p2T7^{Ti}/*Tco*OPB, *T. congolense* p2T7^{Ti}/*Tco*TUB and *T. brucei* p2T7^{Ti}/*Tb*OPB positive clones and induced for RNAi were counted over the 14 day experiment to assess if a growth defect could be detected (Fig. 3.12). Cells were seeded at 1 x 10^6 cells/ml of growth medium. All *T. brucei* parasite clones grew at a higher number of cells/ml when compared to all the *T. congolense* clones. There were no significant differences in the proliferation of both tetracycline induced and non-induced *T. brucei* p2T7^{Ti}/*Tb*OPB parasites where RNAi was successful (Fig. 3.12).

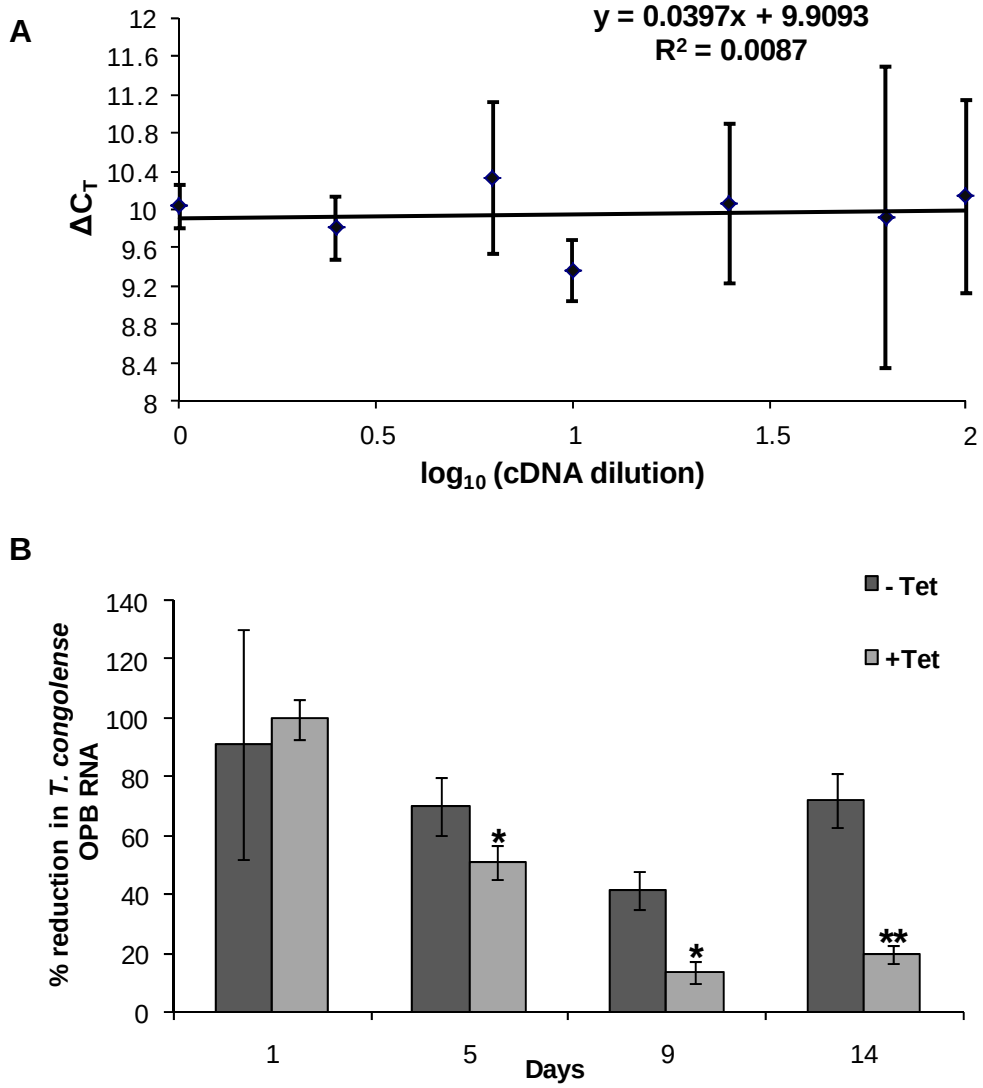


Figure 3.11 qPCR analysis of *T. brucei* p2T7^{Ti}/OPB transformants. A: Validation of the $2^{-\Delta\Delta CT}$ method using actin and OPB primers. Serial dilutions of cDNA were generated from total RNA ranging from 100-0 ng. Average threshold cycle (ΔCt) was calculated ($C_{OPB} - C_{ACT}$) for each cDNA dilution. Error bars were calculated [√S.D (OPB) ^2 + S.D (ACT) ^2] with n = 3. **B: Relative quantification of *Tb*OPB.** *Tb*OPB RNA concentration was calculated using actin as a calibrator in *T. brucei* (*Tb*ACT) on different days in the presence (+Tet) and absence (-Tet) of tetracycline. *Tb*OPB levels relative to *Tb*ACT were plotted as a percentage of mean fold change with day one taken at 100%. Each value is a mean of 3 qPCR reactions. *p< 0.05 and **p<0.001 in comparison to the non-induced control cultures.

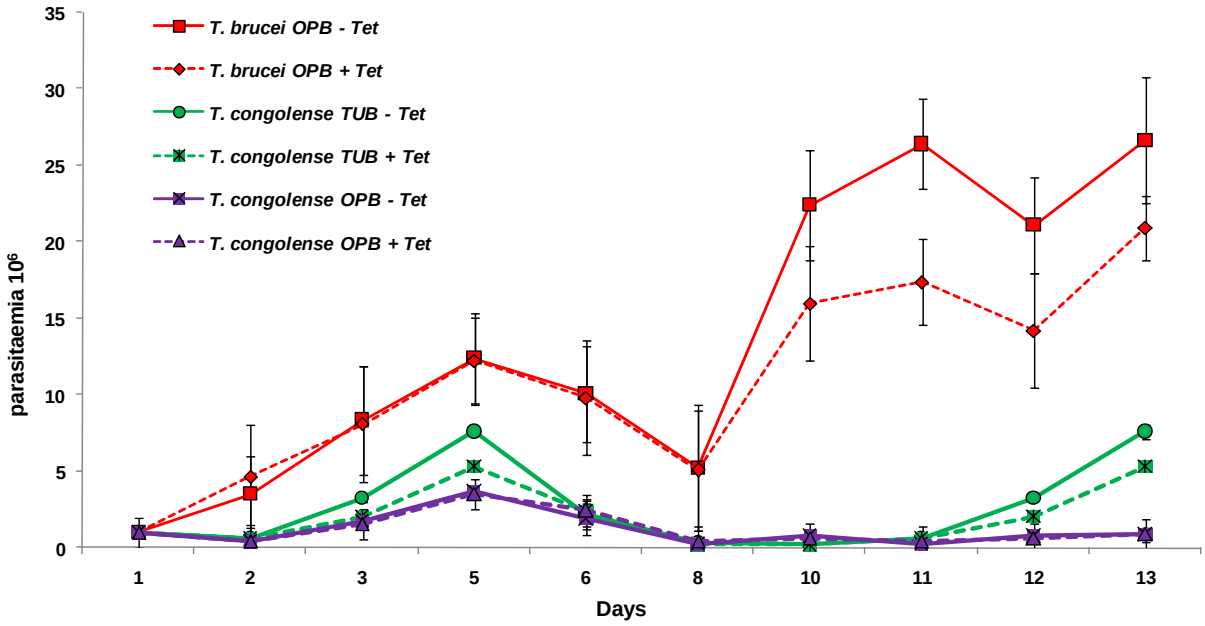


Figure 3.12 Growth curves of transformed parasites after induction with 1 µg/ml tetracycline. Procyclic cultures (5 x 10^5) of each clone: *T. brucei* p2T7^{Ti}/OPB, *T. congolense* p2T7^{Ti}/OPB and *T. congolense* p2T7^{Ti}/TUB, were seeded in duplicate 10 ml flasks and tetracycline either added to induce RNAi (+Tet) or not added in non-induced (-Tet) cultures in a fifteen-day induction experiment. Half the medium in all flasks was replaced on day 8. Numbers of parasites were determined as a mean of three counts ± SEM.

3.4 DISCUSSION

Studies on the *in vivo* function of oligopeptidase B (OPB) in intracellular kinetoplastid parasites led to the description of a variety of roles, including the invasion of non-phagocytic lymphocytes in *T. cruzi* (Caler *et al.*, 1998), the invasion of phagocytic lymphocytes in *L. major* (Munday *et al.*, 2011) and a role for immune evasion by depressing a response from infected macrophages in *L. donovani* (Swenerton *et al.*, 2011). A common result for all these experiments is that there is no phenotype difference between parasites *in vitro* in the absence of OPB except in assays involving cell invasion. Because all these parasites are intracellular kinetoplastids, it was reasoned that OPB may have a different function in *T. brucei* as it exists as an exclusively extracellular parasite (Saravia *et al.*, 2006).

RNAi of *Tco*OPB was unsuccessful as was observed after assessment with anti-OPB peptide antibodies. This failure of RNAi can possibly be attributed to two different factors. More stable RNA transcripts that have a longer half-life are more difficult to degrade than mRNA that is transcribed at lower levels and is more unstable (Bellofatto and Palenchar, 2008). It is also possible for siRNA generated to target genes other than *TcoOPB* (Qiu *et al.*, 2005). This is

especially relevant for the *T. congolense* genome which has not yet been fully sequenced and annotated, leading to a small chance of targeting more than one gene and impaired RNAi. In addition, because RNAi requires siRNA base pairing to targeted RNA, natural variations that occur between species and different variants can lead to ineffectual downregulation of the gene under study (Panjaworayan and Brown, 2011). In chapter 2, the failure of knocking down *Tco*CATL, a multi-copy gene with variations within different catalytic domains (Pillay *et al.*, 2010), creates the possibility for the siRNA to target the different variants rather than the intented peptidase, in this case *Tco*CATL. Variations in the success of RNAi downregulation of genes can be solved by creating RNAi vectors that target multiple sections of one gene, thus increasing the possibility that the intended gene is affected more than other variations (Weinberg and Arbuthnot, 2010).

Down-regulation of *Tb*OPB was successfully achieved using RNAi and this was assessed with anti-OPB peptide antibodies. *Tb*OPB was still detectable in procyclic lysates 14 days after induction using tetracycline, and this could be attributed to *Tb*OPB having a long half-life in culture medium due to the absence of natural inhibitors in mammalian host serum used to prepare complete medium (Coetzer *et al.*, 2008). To confirm if RNAi in transformed *T. brucei* parasites was functional, it was necessary to ascertain if RNA coding for *Tb*OPB has been efficiently degraded as this would provide concrete evidence for the technique (Clayton, 1999). *T. congolense* cultures targeted for tubulin (*Tco*TUB) degradation were used as a control for this experiment and RNAi was induced over a period of three and 14 days. RNAi targeting tubulin has been used as a control because disrupting the gene leads to a visible phenotype where parasites are unable to retain their structure and become “FAT” cells (Inoue *et al.*, 2002).

Although RNAi can be observed in trypanosomes several hours after induction (Coustou *et al.*, 2010), induction using tetracycline was carried out over a period of fourteen days in the present study as this was more effective at reducing levels of *Tco*OPB when compared to *Tco*TUB. Using qPCR, it was possible to detect a down-regulation of *Tb*OPB to 13% of its wild-type level with no observable differences in parasite growth and division. It was concluded that reducing the levels of *Tb*OPB has no deleterious effect on the viability of the parasites and this correlated well with the experiments performed in related kinetoplastid parasites where phenotypes related to the deletion of the *OPB* gene only became apparent

when parasites were evaluated for the invasion of host cells as seen in *T. cruzi* (Caler *et al.*, 1998), *L. major* (Munday *et al.*, 2011) and *L. donovani* (Swenerton *et al.*, 2011).

The definitive functions of different peptidases have been described using RNAi. In *T. brucei*, cathepsin L and B (*Tb*CATL and *Tb*CATB) have been shown with *in vitro* and *in vivo* studies to play a pivotal role in host protein degradation and virulence in mice (Mackey *et al.*, 2004; Abdulla *et al.*, 2008). Targeting cathepsin D for RNAi in *S. mansoni* resulted in the inability of parasites to degrade haemoglobin and led to significant growth retardation *in vitro* and death when used for *in vivo* studies (Morales *et al.*, 2008). In contrast, not all peptidases targeted for RNAi show a phenotype. In contrast to cathepsin D in *S. mansoni*, targeting cathepsin B did not affect the parasites negatively, with results showing the complete degradation of mRNA coding for the peptidase (Skelly *et al.*, 2003). From these experiments, it is clear that RNAi is a first choice for manipulations that lead to the down-regulation of essential genes. If a gene is essential for the survival of an organism, it will be impossible to delete the gene coding for its expression and therefore impossible to study function (Motyka and Englund, 2004; Ullu *et al.*, 2004; Bellofatto and Palenchar, 2008). A major disadvantage of RNAi in trypanosomes is apparent when only trace amounts of an enzyme are required for its function or in cases where an enzyme has a long half-life (Motyka and Englund, 2004). Depending on the organism used for experiments, there is great variability in the length of time it takes for the targeted protein to disappear and in addition, different proteins can differ in the extent to which they can be knocked down by RNAi (Bonetta, 2004). In cases where a protein has a long half-life and is still functional at lower levels of expression, it would be difficult to assign function and necessary to consider either increasing the period of RNAi induction or designing gene knock-out experiments that delete the gene for the target being investigated.

Because *Tb*OPB activity persists after the death of parasites in the host *in vivo* due to the lack of natural inhibitors (Coetzer *et al.*, 2008), it would be difficult to use an *in vivo* model as any new phenotypes could be attributed to residual amounts of the enzyme present before induction. A more attractive model would be the generation of null mutants that would have no OPB available as has been carried out for OPB in *T. cruzi, L. major* and *L. donovani* (Burleigh *et al.*, 1997; Munday *et al.*, 2011; Swenerton *et al.*, 2011). The generation of *T. brucei* OPB knock-out parasites is described in Chapter 4.

CHAPTER 4

GENERATION OF OLIGOPEPTIDASE B NULL MUTANTS IN *TRYPANOSOMA BRUCEI BRUCEI* INCREASES PROLYL OLIGOPEPTIDASE-LIKE ACTIVITY

African trypanosomosis, also referred to as sleeping sickness in humans and nagana in cattle, is a parasitic disease caused by protozoan haemoparasites of the genus *Trypanosoma*, with *T. congolense*, *T. vivax* and *T. brucei brucei* causing nagana (Taylor, 1998). Nagana puts large numbers of cattle in sub-Saharan Africa at risk with an estimated revenue loss of up to five billion US $ annually (Swallow, 2000; McDermott and Coleman, 2001; Antoine-Moussiaux *et al.*, 2009). Symptoms include anaemia, proliferation of lymphoid cells, increased permeability of blood vessels, coagulation and immunosuppression (Vickerman *et al.*, 1993). Many of these symptoms appear as the invading parasites are cleared by the host immune system and most symptoms are attributed to parasite products released by dead and dying parasites (Antoine-Moussiaux *et al.*, 2009). These products include trypanosome lipids, hemolysins, mitogens, inflammatory factors, hepatotoxins and enzymes such as peptidases, phospholipases and acid phosphatises (Tosomba *et al.*, 1996; Taylor and Authié, 2004). This observation has led to the idea of developing vaccines and chemotherapeutic agents that target pathogenic factors released by the parasite during infection, a strategy referred to as an “anti-disease” approach (Authié *et al.*, 2001). Several enzymes have been identified as possible pathogenic factors, with oligopeptidase B in *T. brucei* (*Tb*OPB) and prolyl oligopeptidase (*Tb*POP) implicated in neuroendocrine symptoms during *T. brucei* infection (Knowles *et al.*, 1987; Tetaert *et al.*, 1993; Coetzer *et al.*, 2008; Bastos *et al.*, 2010). Both *Tb*POP and *Tb*OPB belong to the S9 prolyl oligopeptidase family of serine proteases but cleave peptides differently with *Tb*POP hydrolysing substrates at the carboxyl end of prolyl and alanyl residues (Bastos *et al.*, 2010). *Tb*OPB on the other hand only hydrolyses low molecular mass peptides < 30 amino acids that contain dibasic amino acid residues (Arg or Lys) in P_1 with a preference for arginine residues in both P_1 and P_2 (Coetzer *et al.*, 2008).

In trypanosomes, *Tb*OPB is released as an active parasite peptidase by dead and dying parasites during infection (Troeberg *et al.*, 1996a; Morty *et al.*, 2001; Coetzer *et al.*, 2008). The peptidase is expressed by several trypanosome species including the haemoflagellates *T. b. brucei* (Troeberg *et al.*, 1996a), *T. congolense* (Morty *et al.*, 1999a), and *T. vivax* (Coetzer *et al.*, 2008), *T. evansi* (Morty *et al.*, 2005a) which causes surra and the intracellular

T. cruzi (Burleigh *et al.*, 1997) which causes Chagas disease. Due to the restricted size of the substrate binding pocket, plasma protease inhibitors such as serpins and α_2-macroglobulin are not able to inhibit *Tb*OPB in host blood where it remains active (Morty *et al.*, 1999b; Morty *et al.*, 2001; Coetzer *et al.*, 2008). In *T. evansi*, OPB has been implicated in the inactivation of host atrial natriuretic factor (ANF), resulting in an increased blood volume which is associated with many lesions in the circulatory system of trypanosome infected hosts (Ndung'u *et al.*, 1992; Morty *et al.*, 2005a). In *T. cruzi*, an intracellular parasite, OPB is involved in the invasion of mammalian host cells by triggering the release of Ca^{2+} from intracellular stores leading to the recruitment and fusion of host cell lysosomes at the invasion site (Burleigh *et al.*, 1997; Caler *et al.*, 1998). Δ*opb* null-mutant *T. cruzi* parasites were unable to invade host cells due to their inability to mobilise Ca^{2+} (Caler *et al.*, 1998).

Gene knock-out studies in *Leishmania major* and *L. donovani* showed that a loss of OPB resulted in delayed foot pad lesion formation in mice but no difference was observed in virulence during infection (Munday *et al.*, 2011; Swenerton *et al.*, 2011). Knock-out parasites were also less able to infect and survive within macrophages *in vitro* and a large up-regulation of enzymatically inactive membrane associated enolase was detected in *L. donovani* parasites without OPB (Swenerton *et al.*, 2011). Although *T. brucei* does not invade non-phagocytic and phagocytic cells in the mammalian host bloodstream, there is substantial evidence that the parasites leave the blood circulation and invade tissues including crossing the blood brain barrier in humans (Grab *et al.*, 2004; Nikolskaia *et al.*, 2006a; Nikolskaia *et al.*, 2006b; Grab and Kennedy, 2008; Grab *et al.*, 2009), cattle (Losos and Ikede, 1972; Naessens, 2006), mice and rats (Losos and Ikede, 1972; Ojok *et al.*, 2002; Masocha *et al.*, 2006; Naessens, 2006). The *T. b. brucei* Lister 427 strain shows a preferential tropism for mouse testes (Claes *et al.*, 2009) where parasites are not readily cleared by drugs after crossing microvascular endothelial cells into the interstitial tissue between the seminiferous tubules.

Studies into the function of *Tb*OPB during infection have so far been based on experiments using inhibitors, trypanocidal agents and neutralising antibodies (Morty *et al.*, 1998; Morty *et al.*, 1999a; Morty *et al.*, 2000; Morty *et al.*, 2005a; Morty *et al.*, 2005b). To fully understand what role *Tb*OPB plays in parasite physiology and host pathogenesis, oligopeptidase B null mutants (Δ*opb*) were generated in a *T. b. brucei* Lister 427 strain and an attempt was also made at knocking out *OPB* in *T. congolense* IL3000. *Tb*OPB was found not to be necessary

for the survival of trypanosome parasites *in vitro* and during infection in mice with no differences in morphology, rate of parasite division or virulence. Immunohistopathology studies of infected mouse testes also showed that *Tb*OPB is not involved in *T. b. brucei* parasites' ability to cross the blood endothelial barrier. It was, however, observed that *Tb*POP-like activity was significantly increased in Δ*opb* null mutant parasites with no associated increase in *Tb*POP protein levels. These data may suggest that another POP-like peptidase was responsible for the observed increase in activity or that *Tb*OPB is responsible for generating an endogenous TbPOP inhibitor.

4.1 MATERIALS AND METHODS

4.1.2 Materials

Molecular biology: ApaI, NotI, XbaI, shrimp alkaline phosphatase (SAP), T4 DNA ligase, 10 mM dNTP mix, X-gal, IPTG, MassRuler™ DNA ladder mix, GeneJET™ plasmid miniprep kit, TransformAid™ bacterial transformation kit, shrimp alkaline phosphatase (SAP) and Biotin DecaLabel™ DNA labeling kit were obtained from Fermentas (Vilnius, Lithuania). The pGEM-T® vector was obtained from Promega (Madison, WI, USA), DNA clean and concentrator kit from ZymoResearch (Orange, CA, USA), E.Z.N.A® gel extraction kit from PEQlab (Erlangen, Germany) and *Escherichia coli* JM 109 cells were purchased from New England Biolabs (Ipswitch, MA, USA). FIREpol® Taq polymerase, 10 x PCR reaction buffer and 25 mM $MgCl_2$ were acquired from Solis Biodyne (Tartu, Estonia). Hybond™-N DNA binding membrane was obtained from Amersham (Buckinghamshire, UK). Molecular biology grade agarose was purchased from Conda laboratories (Madrid, Spain), Ampicillin sodium salt from USB Corporation (Cleveland, OH, USA), bacteriological agar and tryptone from Merck (Darmstadt, Germany). The yeast extract was purchased from Sigma (St. Louis, MO, USA).

Trypanosome culture: Minimum essential medium (MEM), Iscove's Modified Dulbecco's Medium (IMDM), HEPES, sodium pyruvate, thymidine, adenosine, hypoxanthine, thymidine, bathocuprone sulfate, glutamine, hemin, proline, cis-aconitate, tetracycline, doxycycline hyclate and cyclophosphamide were obtained from Sigma (St. Louis, MO, USA). Heat inactivated goat serum and heat inactivated foetal calf serum (FCS) were obtained from Gibco (Paisley, UK). 0.2 µm filters were purchased from Pall (Ann Arbor, USA). 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 H solution 2 were

obtained from Lonza (Levallois-Perret, France). *T. congolense* TRUM183:29-13 and *T. brucei* T7T-29 RNAi inducible parasite strains and NOD/SCID mice were a gift from Professor Theo Baltz (University of Victor-Segalen, Bordeaux 2, France). Balb/C mice were obtained from the Bioresources Unit (BRU), Westville campus, University of KwaZulu-Natal (UKZN). Protocols using mice were approved by the UKZN animal ethics committee (Reference 036/09/Animal).

Immunofluorescence, western blotting and microscopy: Fluorescein (FITC) conjugated donkey anti-mouse IgG, Cy3 (water soluble cyanine fluorescent dye) conjugated donkey anti-chicken IgG and horseradish peroxidase conjugated horse anti-mouse IgG were obtained from Vector labs (Burlingame, USA). Goat anti-rabbit Dylite™ conjugate was purchased from Jackson labs (Baltimore, USA). Rabbit anti-VSG serum was a kind gift from Dr Philipe. Büscher, Tropical veterinary institute, Antwerp, Belgium. The mouse anti-POP serum for *Tb*POP was from Dr. Aime Santana, Laboratório de Interção Parasito-Hospedeiro, Faculdade de Medicina, Universidade Brasília city, Brazil and Dr. Philippe Grellier, Muséum National d'Histoire Naturelle, Paris, France. Mouse anti-OPB serum was available from a previous study (Bizaaré, 2008). The LSM 710 confocal microscope was purchased from Zeiss, Göttingen, Germany.

Peptide substrate: The peptide substrates Benzyloxycarbonyl (Z)-Phe-Arg-7-amino-4-methylcoumarin (AMC), Z-Arg-Arg-AMC, Z-Gly-Pro-AMC and Z-Arg-Gly-Phe-Phe-Pro-4MβNA·HCl were obtained from Bachem (Torrance, CA, USA). Activity assays were measured using a FLUOstar optima spectrophotometer from BMG Labtech (Offenburg, Germany).

4.2.2 Methods

4.2.2.1 Trypanosome culture

All blood-stream forms of *T. b. brucei* Lister 427 wild-type and knock-out clones generated were cultured in supplemented Iscove's Modified Dulbecco's Medium (IMDM) as previously described (Hirumi and Hirumi, 1989). Briefly, 16.66 g of IMDM was dissolved in 1 litre of sterile water containing 3.6 mM $NaHCO_3$, 1 mM hypoxanthine, 1 mM sodium pyruvate, 0.16 mM thymidine, 0.05 mM bathocuprone, 1.5 mM L-cysteine and 0.2 mM 2-mercaptoethanol. The pH was adjusted to between 7.2 and 7.4 and 10% (v/v) heat inactivated foetal calf serum added before filtration using a 0.2 µm filter. *T. congolense* IL3000 procyclic cultures were maintained as previously described in Section 2.2.2.2 but without hygromycin or neomycin.

4.2.2.2 Generation of heterozygote Δ*opb* null mutants in *T. brucei* and *T. congolense*

The sequences flanking the *OPB* gene in *T. b. brucei* Lister 427 (*TbOPB*) and *T. congolense* IL3000 (*TcoOPB*) were identified in the GeneDB database (http://www.genedb.org/tbrucei/, and http://www.genedb.org/tcongolense/, accessed 10-10-2007) and used to design primers that flank the 3' and 5' regions of the *OPB* gene in *T. brucei* and *T. congolense* (Table 4.1). Genomic DNA was extracted from *in vitro* cultures of *T. b. brucei* Lister 427 and *T. congolense* IL3000 as previously described (Medina-Acosta and Cross, 1993) and used as a template for PCR. T-vector cloning and sub-cloning into knock-out vectors was carried out as previously described in Section 2.2.2.5, resulting in four recombinant knock-out vectors: pGLbla*Tb*OPB and pGLneo*Tb*OPB for *T. brucei* blasticidin and neomycin resistance plasmids respectively, and pGLbla*Tco*OPB and pGLneo*Tco*OPB for *T. congolense*.

T. b. brucei Lister 427 bloodstream parasites and *T. congolense* IL3000 procyclic parasites were transfected with recombinant knock-out constructs using the Amaxa Nucleofector system as previously described in Section 2.2.2.2 with appropriate drugs for selection (Burkard *et al.*, 2007). Briefly, a pellet of 10^7 parasites was resuspended in 100 µl Basic Parasite Nucleofector H solution 2, mixed with 10 µg of NotI linearised pGLbla*Tb*OPB for *T. brucei* parasites and pGLbla*Tco*OPB for *T. congolense* parasites before transfection using the X-001 program setting on the Amaxa transfection system (Lonza, Basel, Switzerland) (Burkard *et al.*, 2007). Stably transfected trypanosomes were first cultured for 24 h in either supplemented IMDM medium (*T. brucei*) or complete MEM (*T. congolense*) before selection with blasticidin (5 µg/ml). Single knock-out clones, generated by limiting dilution, were expanded and taken through a second round of transfection with NotI linearised pGLneo*Tb*OPB or pGLbla*Tco*OPB and selected with neomycin (2.5 µg/ml) and blasticidin (5 µg/ml) as described above.

4.2.2.3 Confirmation of clones using Southern blot with digoxigenin labelled probes

To confirm the stable transfection of both plasmids and the deletion of the *OPB* gene in *T. b. brucei* Lister 427 strain and *T. congolense* IL3000, a Southern blot was carried out using digoxigenin probes according to the manufacturer's protocol. For *T. brucei*, two probes that spanned the 3' flanking region of the *T. b. brucei OPB* gene and either the blasticidin or the neomycin coding regions were prepared by PCR incorporating digoxigenin labelled nucleotides. A third probe that spanned the *T. b. brucei* OPB gene was also prepared. For *T. congolense*, a probe that recognised the blasticidin gene was prepared. *T. brucei* DraI and *T.*

congolense BamHI restricted genomic DNA isolated from the selected clones was separated on a 1% (w/v) agarose gel at 110V for 4 h, transferred onto a DNA binding nylon membrane using capillary action before probing with the digoxigenin probes prepared as described above.

Table 4.1 Sequences of primers designed to amplify the 5' and 3' flanking regions of oligopeptidase B (*OPB*) from *T. congolense* strain IL3000 and *T. brucei* 427 Lister for cloning into the knock-out vectors pGLneo and pGLbla. Each insert primer had either a NotI, XbaI or ApaI restriction site (underlined) for incorporation into the pGLneo or pGLbla knock-out plasmids. Neomycin and blasticidin resistance gene primers were used to confirm orientation of the 3' flanking regions in the final recombinant plasmids

Primer name and restriction site	Sequence (5'→3', restriction site underlined)
OPB *T. congolense* 3' flank reverse (ApaI)	CGGGGCCCCTTTCACGATCAAAGAAATAAATCT
OPB *T. congolense* 3' flank forward (ApaI)	CGGGGCCCGCGCAAGTGGGGGAGGGGCGGTGAGA
OPB *T. congolense* 5' flank reverse (XbaI)	CGTCTAGATGTGCGCGTGCGTGTAATAAGCTTCGCCCC
OPB *T. congolense* 5' flank forward (NotI)	CGGCGGCCGCCCGACTACTTTCAGCGTCGAAAGCGATTGA
OPB *T. brucei* 5' flank reverse (XbaI)	CGTCTAGAGATTGCGATGGAAAGTGCCTCCGAG
OPB *T. brucei* 5' flank forward (NotI)	CGGCGGCCGCCCGAATACGTGCAGGTGACAATATA
OPB *T. brucei* 3' flank reverse (ApaI)	CGGGGCCCGGTCACTCACATGCAACACGCGGCAA
OPB *T. brucei* 3' flank forward (ApaI)	CGGGGCCCGTAGGTAAGGAATAGCTGAAACTTGT
pGLneo forward primer (Neo Fw)	CGTGTTCCGGCTGTCAGCGC
pGLneo reverse primer (Neo Rv)	ACGACGAGATCCTCGCCGTC
pGLbla forward primer (Bla Fw)	TCAACAGCATCCCCATCTCT
pGLbla reverse primer (Bla Rv)	GCAGCAATTCACGAATCCCA

4.2.2.4 Western blot analysis of *T. brucei* Δ*opb* clones

A pellet of 2 x 10^6 parasites from each culture was washed twice with PBS (pH 7.2) and resuspended in 50 µl of lysis buffer [10 mM EDTA, 1% (v/v) Triton X-100, 20 mM Tris-HCl buffer, pH 7.2, containing 1 mM 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF) and 10 µM L-trans-epoxysuccinyl-leucylamido (4-guanidino) butane (E64)], an equal volume of reducing gel loading buffer was added and samples boiled for 10 min. Lysates generated (25 µl) from ~ 5 x 10^5 parasites were separated per well on a 10% SDS-PAGE gel

at 18 mA per gel, blotted onto nitrocellulose (Towbin *et al.*, 1979) and probed with appropriate antibodies. The primary antibodies used were polyclonal mouse anti-OPB (1:200) and polyclonal chicken anti-immunoglobulin binding protein (BIP) (1:1000) and mouse anti-POP serum for *Tb*POP (1:300). Horseradish peroxidise conjugated horse anti-mouse IgG and rabbit anti-chicken IgG secondary antibodies were used at 1:10000. The blot was developed using 0.06% (w/v) 4-chloro-1-naphthol, 0.0015% (v/v) H_2O_2 in Tris buffered saline (20 mM Tris-HCl, 200 mM NaCl, pH 7.4).

4.2.2.5 Enzymatic activity analysis of *T. b. brucei* Δ*opb* clones

Parasites (1 x 10^7), were removed from culture, washed twice in PBS (pH 7.2) and resuspended in 900 µl of 0.1% (w/v) Brij-35 containing 10 µg/ml soybean trypsin inhibitor (SBTI) to inhibit other parasite serine proteases, 10 µM E-64 and 1 mM EDTA. The lysate was incubated for 10 min on ice, 100 µl of 10 x PBS added and centrifuged (10 000 *g*, 5 min, 4°C). The protein concentration of the lysate supernatant was determined using the BCA™ protein assay. The hydrolysis of Z-Arg-Arg-AMC, Z-Gly-Pro-AMC and Z-Arg-Gly-Phe-Phe-Pro-4MβNA.HCl by the total parasite protein extract (5 µg) was determined for each clone. Briefly, parasite protein extract diluted in 0.1% (w/v) Brij-35 was incubated with assay buffer [200 mM Tris-HCl buffer, pH 8, 10 mM DTT and 0.02% (w/v) NaN_3] for 10 min at 37°C. Aliquots were combined with the appropriate AMC substrate (20 µM) and the fluorescence was read (excitation at 360 nm and emission at 460 nm) using a FLUOstar OPTIMA fluorescence microplate reader. Lysis buffer [0.1% (w/v) Brij-35, containing 10 µM E-64, 10 µg/ml SBTI, 1 mM EDTA and 1 x PBS] was used as a negative control. Parallel activity assays with inhibitors were carried out with 1 mM AEBSF for *Tb*OPB, and 100 µM Tosyl phenylalanyl chloromethyl ketone (TPCK) for *Tb*POP. When determining cysteine peptidase levels in total parasite extract, hydrolysis of Z-Phe-Arg-AMC was carried out using the same protocol but with 10 µM E-64 replaced by 1 mM AEBSF and 1 µg/ml of pepstatin A. For statistical analyses, values were expressed as means ± standard error of the mean (SEM). Significance levels were calculated by unpaired t tests and differences were considered significant at a p value <0.001.

4.2.2.6 Detection of enzyme activity in zymograms

Parasites (2 x 10^6) were removed from culture, washed twice with PBS (pH 7.2) and resuspended in 50 µl of lysis buffer [10 mM EDTA, 1% (v/v) Triton X-100, 20 mM Tris-HCl

buffer, pH 7.2]. An equal volume of non-reducing gel loading buffer was added. Lysates generated (25 µl) from ~ 5 x 10^5 parasites were separated per well on a 10% SDS-PAGE gel co-polymerised with 1% (w/v) gelatin at 18 mA per gel. After electrophoresis, gels were washed twice for 30 min in 2.5% (v/v) Triton X-100 before incubation in 0.1 M Mes buffer pH 6.0, containing 2 mM EDTA and 1 mM dithiothreitol (DTT) overnight at 37ºC with different inhibitors. The following inhibitors and concentrations were used: 10 µM E-64, 10 µM pepstatin A, 100 µM TPCK, 100 µg/ml SBTI and 1 mM AEBSF. A control without any inhibitor was also incubated overnight. The gels were stained in freshly prepared 0.1% (m/v) amido black in methanol/ acetic acid/ water (30:10:60, by vol.) and destained in methanol/ acetic acid/ water (30:10:60, by vol.).

4.2.2.7 Immunofluorescence, immuno-histochemistry and microscopy

For immunofluorescence, Δ*opb* and wild type *T. b. brucei* Lister 427 parasites (1 x 10^6) were removed from culture, washed twice in PBS, spread and dried over microscope slides coated with 0.01% (w/v) poly-L-lysine before fixing with methanol for 15 min. The fixed parasite slides were blocked with 1% (w/v) BSA-PBS for 1 h before incubation with mouse anti-OPB serum (1:500) and chicken anti-BIP IgY (5 µg/ml) in 0.2% (v/v) Triton X-100 in PBS for 1 h. All subsequent washes were done in 0.2% (v/v) Triton X-100 in PBS with 5 min for each wash. The slides were washed four times and incubated with the secondary fluorescein (FITC) conjugated donkey anti-mouse IgG and Cy3 (water soluble cyanine fluorescent dye) conjugated donkey anti-chicken IgG both at 1:1000 in 0.2% (v/v) Triton X-100 in PBS for 1 h. The slides were washed six times before incubation with Hoechst stain in PBS (2 ng/ml) for 5 min. The slides were given 5 final washes in PBS, mounted and viewed using a LSM 710 confocal microscope. To examine the ability of *T. b. brucei* Δ*opb* clones to cross endothelial barriers, testes from mice infected with 1 x 10^6 *T. b. brucei* Δ*opb* for 5 days were dissected and 14 µm thick sections prepared using a cryostat. Sections were mounted, fixed and immunostained as previously described (Masocha *et al.*, 2004). Rabbit anti-VSG serum (1:1000) was incubated with tissue sections to detect trypanosomes in tissue and goat anti-rabbit Dylite™ conjugate (1:800) used as a detection antibody. Sections were incubated with Hoechst stain in PBS (2 ng/ml) before viewing using a LSM 710 confocal microscope.

4.2.2.8 Mouse infections

In order to observe what role *Tb*OPB plays in parasite virulence and infection, groups of 5 female BALB/c mice were infected by an intraperitoneal injection of 1 x 10^3 *T. b. brucei* Lister

427 Δ*opb* parasites per mouse resuspended in 50 µl of phosphate saline glucose (PSG) [57 mM Na_2HPO_4, 3 mM NaH_2PO_4, 42 mM NaCl, 50 mM glucose, 1 mM hypoxanthine, pH 7.4] in 50% (v/v) glycerol. A control group was also infected with 1 x 10^3 *T. b. brucei* Lister 427 wild type parasites resuspended in 50 µl of PSG. Parasitaemia was measured daily by bleeding from the tail and survival of mice monitored during infection. Infection and care of infected mice was carried out using protocols approved by the University of KwaZulu-Natal animal ethics committee (Reference 036/09/Animal).

4.3 RESULTS

4.3.1 Generation of *T. b. brucei* Lister 427 and *T. congolense* IL3000 *OPB* mutants.

Primers designed were used to amplify the 626 bp 3' and the 929 bp 5' regions flanking the *T. congolense OPB* gene (*TcoOPB*) (Fig.4.1, A1, lane 1 and B1, lane 1). The 626 bp and 929 bp PCR products amplified from *T. congolense* IL3000 genomic DNA were successfully cloned into the pTZ57R/T vector (T-vector) (Fig. 4.1, A2, lanes 1-5 and B2, lanes 1-8).

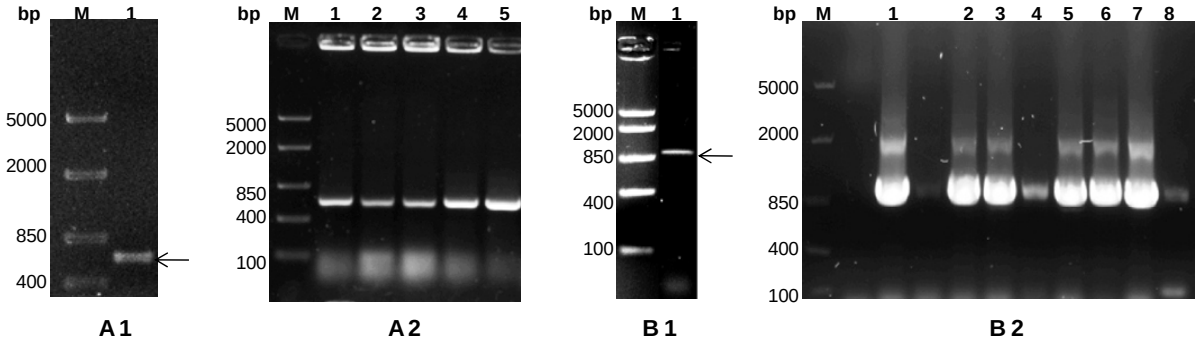


Figure 4.1 Amplification and T-vector cloning of the 5' and 3' regions flanking the *T. congolense* oligopeptidase B gene (*TcoOPB*). M: Middle Range™ DNA ladder mix. **A1: lane 1**, 626 bp PCR product flanking the 3' region of the *TcoOPB* gene (←). **A2: lanes 1-5**, colony PCR identifying positive 3' *TcoOPB* clones in the T-vector. **B1: lane 1**, 929 bp PCR product flanking the 5' region of the *TcoOPB* gene (←). **B2: lanes 2-10**, colony PCR identifying 5' *TcoOPB* positive clones in the T-vector. Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml)**.**

Primers were also used to amplify the 902 bp 5' and the 601 bp 3' regions flanking the *TbOPB* gene in *T. brucei* (Fig. 4.2, A1, lane 1 and B1, lane 1) and subsequently cloned into the T-vector (Fig. 4.2, A2, lane 2 and B2, lanes 1-4).

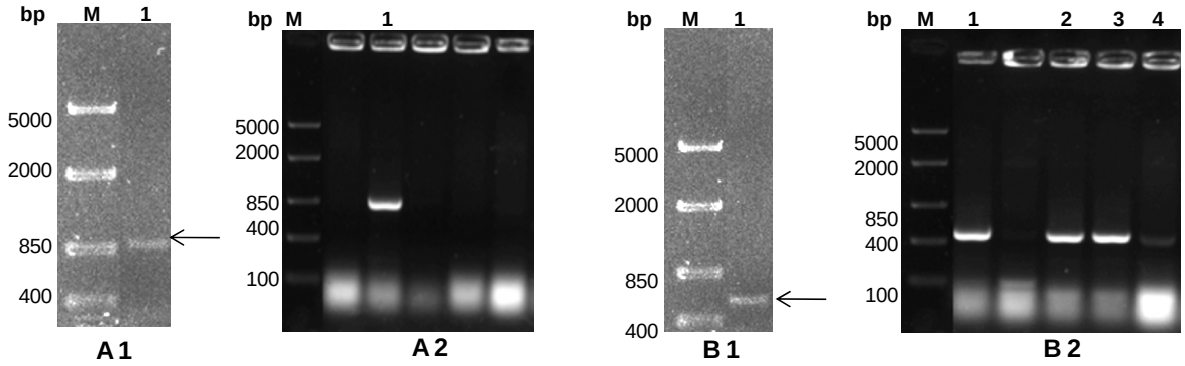


Figure 4.2 Amplification and T-vector cloning of the 5' and 3' regions flanking the *T. brucei* oligopeptidase B gene (*TbOPB*). M: Middle Range™ DNA ladder mix. **A1: lane 1**, 902 bp PCR product flanking the 5' region of the *TcoOPB* gene. **A2: lane 1**, colony PCR identifying a 5' *TbOPB* positive clone positive in the T-vector **B1: lane 1**, 601 bp PCR product flanking the 3' region of the *TcoOPB* gene. **B2: lanes 1-4**, colony PCR identifying 3' *TbOPB* positive clones in the T-vector. Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml)**.**

Vectors containing blasticidin resistance (pGLbla) and neomycin resistance (pGLneo) were prepared for ligation with 3'- and 5'- *TcoOPB,* or *TbOPB*. pGLbla (4566 bp) (Fig. 4.3, A1, lane 1) and pGLneo (4995 bp) (Fig. 4.3, A1, lane 2) were restricted with either NotI/XbaI or ApaI. Recombinant T-vector clones bearing *T. congolense* sequences were restricted using enzymes originally incorporated into the primers (Table 4.1), releasing the 626 bp product flanking the 3' region (Fig. 4.3, A2, lane 1) and the 929 bp product flanking the 5' region of the *TcoOPB* gene (Fig. 4.3, A3, lane 1). The two restricted products were sub-cloned into pGLneo and pGLbla to give pGLneo*TcoOPB* and pGLbla*TcoOPB* respectively. Orientation of the 3' insert (due to identical ApaI sites on both ends of the restricted product) in the pGLbla*TcoOPB* and pGLneo*TcoOPB* recombinant plasmids was confirmed by PCR using vector and insert primers resulting in a 1843 bp product for pGLbla*TcoOPB* (Fig. 4.3, B1, lane 1) and a 1893 bp product for pGLneo*TcoOPB* (Fig. 4.3, C, lane 1). The 929 bp 5' and 626 bp 3' inserts in pGLbla*TcoOPB* (Fig. 4.3, B2, lanes 1 and 2) and in pGLneo*TcoOPB* (Fig. 4.3 C lanes 2 and 3), together with the 314 bp blasticidin (Fig. 4.3, B2, lane 3) and the 435 bp neomycin resistance markers (Fig. 4.3, C, lane 4) were also validated using PCR.

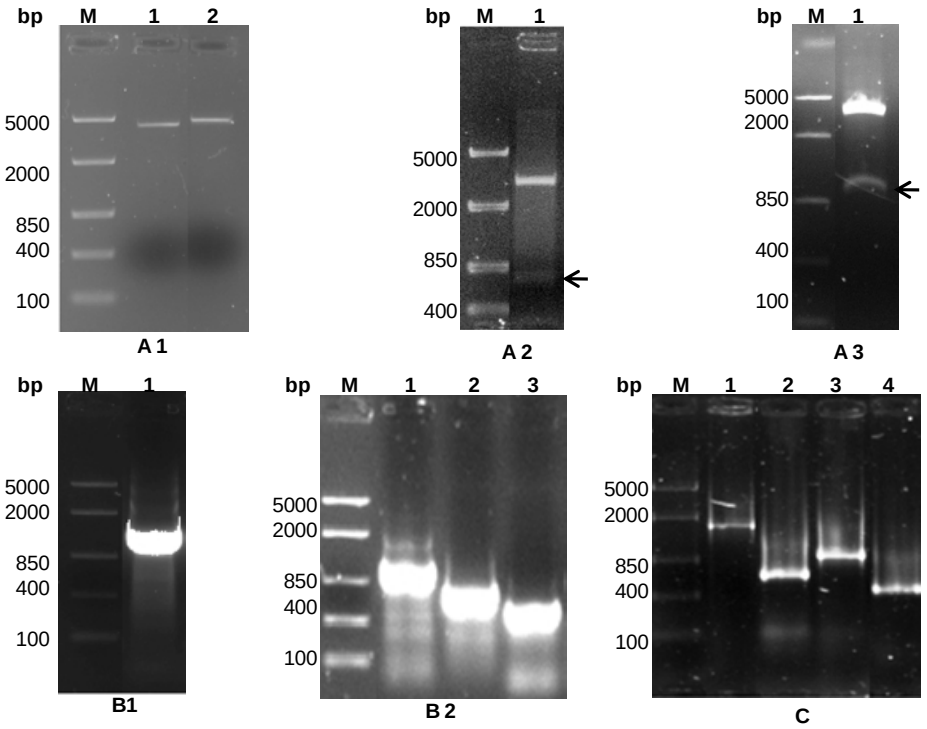


Figure 4.3 Subcloning of the 5' and 3' regions flanking the *TcoOPB* gene into pGLbla and pGLneo knock-out vectors. M: Middle Range™ DNA ladder mix. **A1:** Vector preparation; **lane 1**, NotI/XbaI or ApaI restricted 4566 bp pGLbla; **lane 2**, NotI/XbaI or ApaI restricted 4995bp pGLneo. **A2: lane 1**, NotI/XbaI restricted 626 bp product flanking the 3' region of the *TcoOPB* gene (←). **A3: lane 1**, **ApaI** restricted 929 bp product flanking the 5' region of the *TcoOPB* gene (←). **B1: lane 1**, 1843 bp product from recombinant pGLbla confirmed with PCR using the blasticidin forward primer with the insert 3' reverse primer. **B2: lane 1**, 929 bp product from recombinant pGLbla confirmed with PCR using 5' insert primers; **lane 2**, 626 bp product from recombinant pGLbla confirmed with PCR using 3' insert primers; **lane 3**, 314 bp product from recombinant pGLbla confirmed using blasticidin forward and reverse primers. **C: lane 1**, 1893 bp product from recombinant pGLneo confirmed with PCR using the neomycin forward primer with the insert 3' reverse primer; **lane 2**, 626 bp product from recombinant pGLneo confirmed with PCR using 3' insert primers; **lane 3**, 929 bp product from recombinant pGLneo confirmed with PCR using 5' insert primers; **lane 4**, 435 bp product from recombinant pGLneo confirmed with PCR using neomycin forward and reverse primers. Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml)**.**

Recombinant *T. brucei* T-vector clones were also restricted releasing the 902 bp 5' (Fig. 4.4, A1, lanes 1 and 2), and the 601 bp 3' (Fig. 4.4, A2, lanes 1-3) regions flanking the *TbOPB* gene. The two restricted products were sub-cloned into pGLneo and pGLbla (Fig. 4.3, A1, lane 1 and 2) to give pGLneo*TbOPB* and pGLbla*TbOPB* respectively. Orientation of the 3' insert was confirmed by PCR resulting in a 2007 bp product for pGLneo*TbOPB* (Fig. 4.4, B, lane 3) and a 1868 bp product for pGLbla*TbOPB* (Fig. 4.4, C1, lane 1). The 601 bp 3' inserts (Fig. 4.4, B, lane 2 and C2, lane 3) and the 902 bp 5' inserts (Fig. 4.4, B, lane 1 and C2, lane 1) together with the 435 bp neomycin and 314 bp blasticidin resistance markers (Fig. 4.4, B, lane 4 and C2, lane 2) were also validated using PCR.

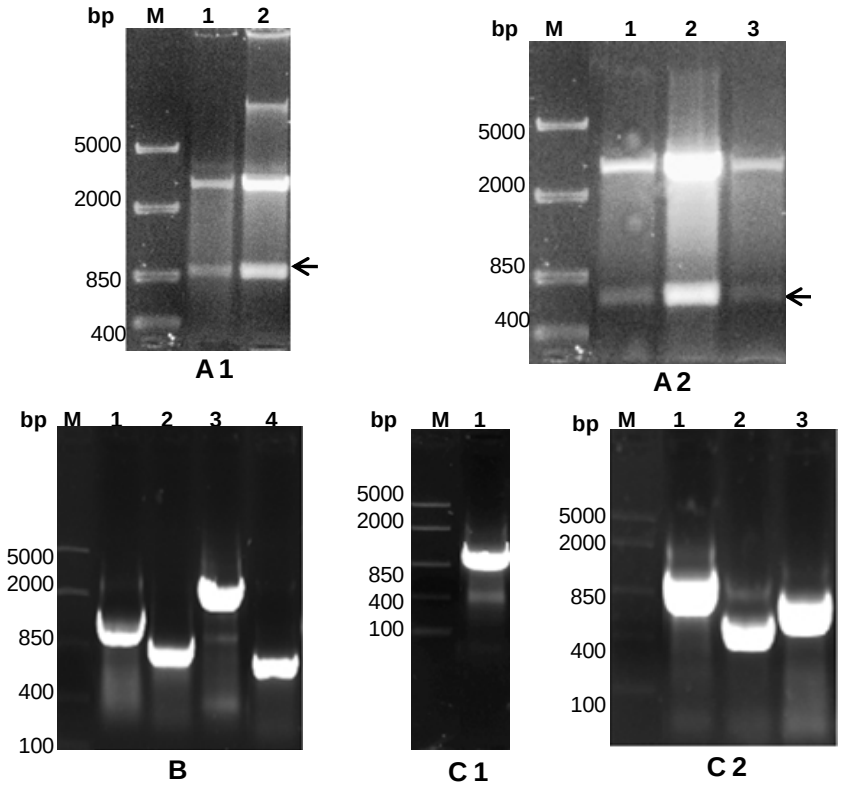


Figure 4.4 Subcloning of the 5' and 3' regions flanking the *TbOPB* gene into pGLbla and pGLneo knock-out vectors. M: Middle Range™ DNA ladder mix. **A1: lanes 1-2**, NotI/XbaI restricted 902 bp product flanking the 5' region of the *TbOPB* gene (←). **A2: lanes 1-3**, ApaI restricted 601 bp product flanking the 3' region of the *TbOPB* gene (←). **B: lane 1**, 902 bp product from recombinant pGLneo confirmed with PCR using 5' insert primers; **lane 2**, 601 bp product from recombinant pGLneo confirmed with PCR using 3' insert primers; **lane 3**, 2007 bp product from recombinant pGLneo confirmed with PCR using neomycin forward primer with insert 3' reverse primer; **lane 4**, 453 bp product from recombinant pGLneo confirmed with PCR using neomycin forward and reverse primers. **C1: lane 1**, 1868 bp product from recombinant pGLbla confirmed with PCR using blasticidin forward primer with insert 3' reverse primer. **C2: lane 1**, 902 bp product from recombinant pGLbla confirmed with PCR using 5' insert primers; **lane 2**, 314 bp product from recombinant pGLbla confirmed using blasticidin forward and reverse primers; **lane 3**, 626 bp product from recombinant pGLbla confirmed with PCR using 3' insert primers. Products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (3 µg/ml)**.**

The transfection of wild type bloodstream forms of *T. congolense* with pGLbla*Tco*OPB produced a distinct heterozygote clone resistant to blasticidin (BLA), *T. congolense Δopb*::*BLA/OPB*. Several rounds of transfection with pGLneo*TcoOPB* using the *T. congolense Δopb*::*BLA/OPB* were not successful (results not shown). The *T. congolense* heterozygote clone was confirmed in a Southern blot using a probe specific to blasticidin (Fig. 4.5 A). A band at the expected size of 2470 bp was detected in BamHI restricted *T. congolense* DNA *Δopb*::*BLA/OPB* (Fig. 4.5 B).

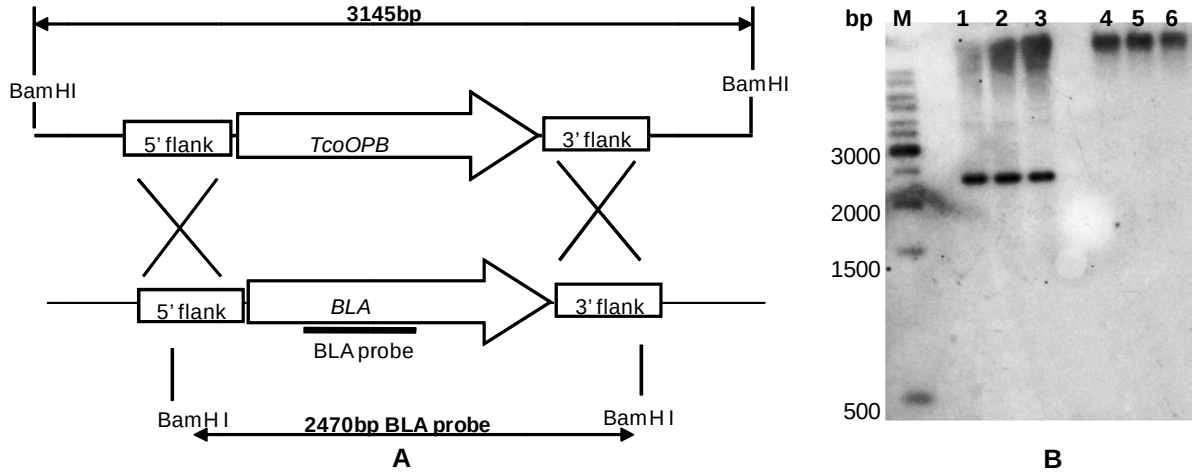


Figure 4.5 Identification of a *Tco*OPB heterozygote clone using a blasticidin probe. A: Schematic diagram showing the *TcoOPB* locus. *TcoOPB* and blasticidin (*BLA*) genes are shown by ⇨, 5' and 3' flanking regions by ▭. The blasticidin probe prepared and the size expected on Southern blot after restriction with BamHI is shown in bold. **B:** Southern blot of *T. congolense Δopb*::BLA/OPB heterozygote probed with BLA. **M:** MassRuler™ DNA ladder mix; **lanes 1 - 3**, *Tco*OPB heterozygote genomic DNA; **lanes 4 - 6**, wild-type *T. congolense* IL3000 strain genomic DNA.

The transfection of wild type bloodstream forms of *T. b. brucei* with pGLbla*TbOPB* produced a distinct heterozygote clone resistant to blasticidin (BLA), ∆*opb*::*BLA/OPB*. A parallel transfection of wild type bloodstream forms of *T. b. brucei* with pGLneo*TbOPB* yielded a second heterozygote clone resistant to neomycin (NEO), ∆*opb*::*NEO/OPB*. A second round of transfection with pGLneo*TbOPB* using single knock-out ∆*opb::BLA/OPB* resulted in the selection of a null mutant clone, *T. b. brucei* Lister 427 ∆*opb* null mutant, resistant to both blasticidin and neomycin. The deletion of *TbOPB* was confirmed by Southern blot using three probes that spanned the *OPB* gene [Fig. 4.6, probe 1 (*OPB*)], a section of the *TbOPB* 3' flanking region with part of the gene coding for neomycin [Fig. 4.6, probe 2 (*NEO*)] and a section of the *TbOPB* 3' flanking region with part of the gene coding for blasticidin [Fig. 4.6, probe 3 (*BLA*)].

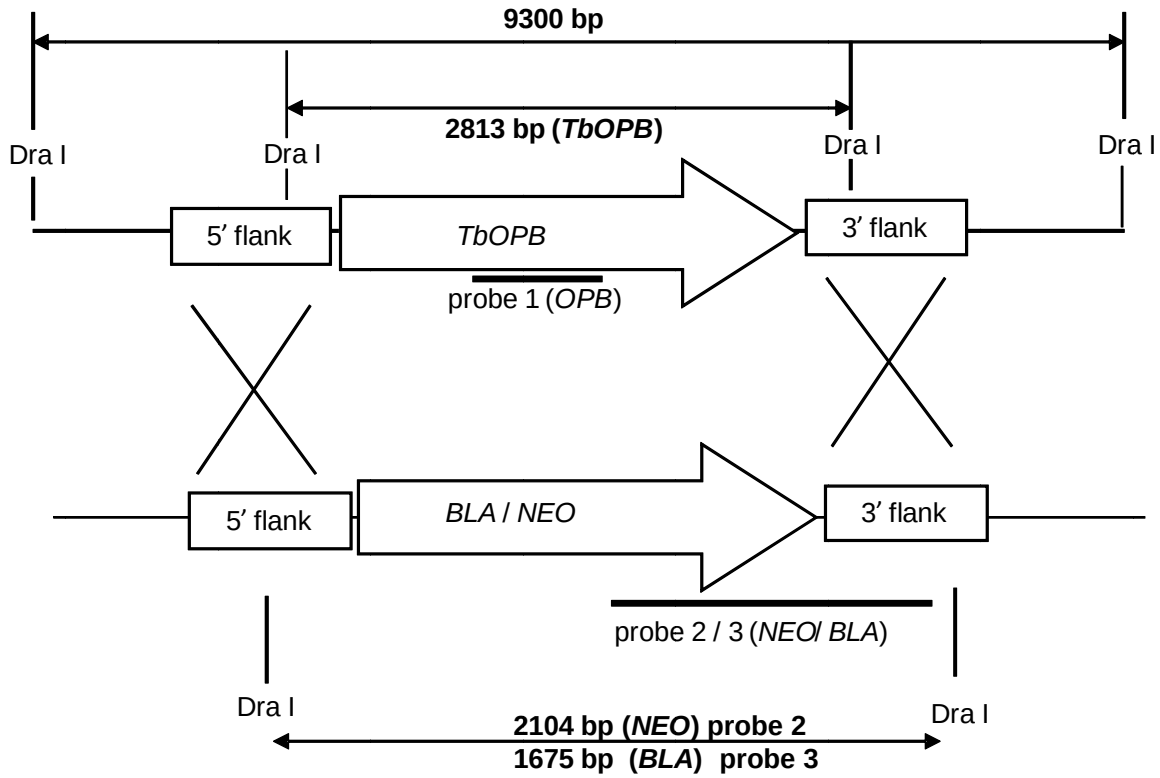


Figure 4.6 Preparation of probes spanning the *OPB* gene in *T. b. brucei* Lister 427. Schematic diagram showing the *TbOPB* locus. *TbOPB*, blasticidin (*BLA*) and neomycin (*NEO*) genes are shown by ⇨ , 5' and 3' flanking regions by ▭. Probes prepared and the sizes expected on a Southern blot after restriction with DraI are in bold.

A Southern blot using probe 1 for *TbOPB* revealed a 2813 bp fragment in DraI restricted wild type *T. b. brucei* Lister 427 genomic DNA containing full length *TbOPB* (Fig. 4.7 A). No band was detected when the same probe was used for DraI restricted Δ*opb* null mutant genomic DNA (Fig. 4.7 B). To ensure that both resistance genes were in the right locus, a Southern blot using DraI restricted Δ*opb* null mutant parasite genomic DNA with probes 2 and 3 revealed a 2104 bp fragment for *NEO* (Fig. 4.7 C and D) and a 1675 bp fragment for *BLA* (Fig. 4.7 C and D). Probes 2 and 3 each recognised both resistance gene sizes as they had a section of the 3' flanking region in common.

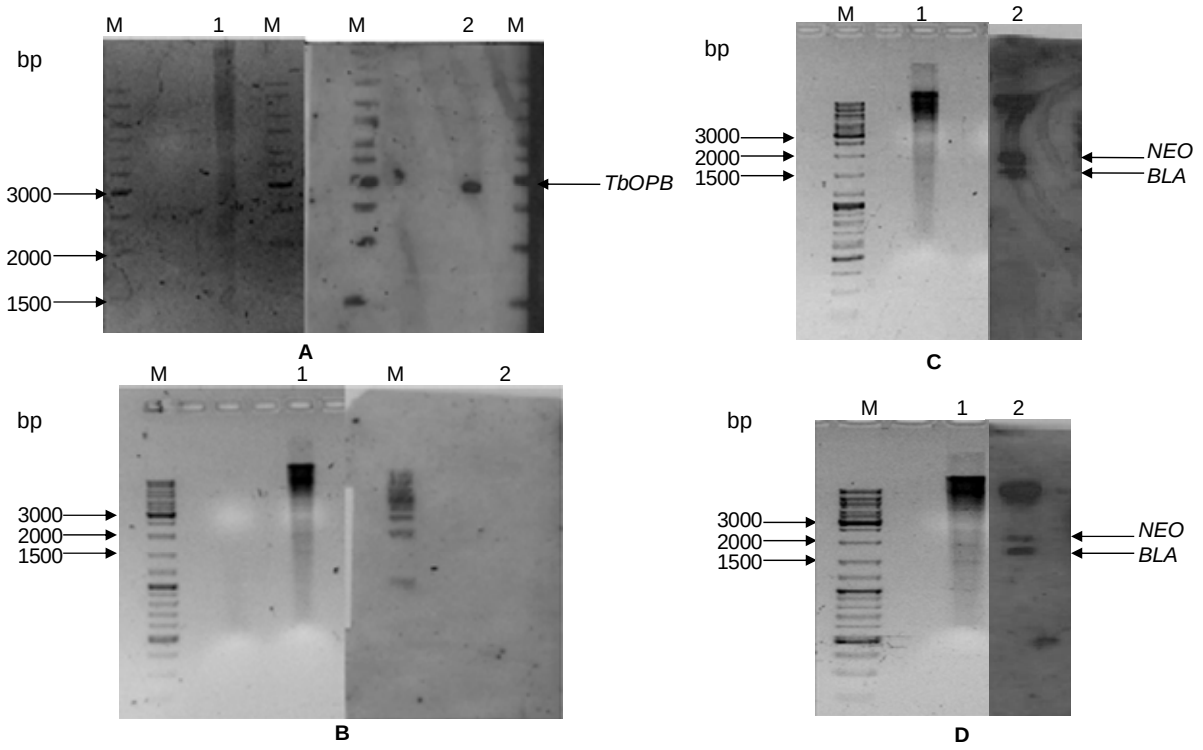


Figure 4.7 Deletion of the OPB gene in *T. b. brucei* Lister 427. Agarose gel (left hand panels) and corresponding Southern blot (right hand panels) of wild type and ∆*opb* null mutants. **M**: DNA ladder; **lane 1**, DraI restricted genomic DNA; **lane 2**, corresponding Southern blot; NEO (neomycin); BLA (blasticidin). **A:** *T. brucei* 427 Lister wild type genomic DNA incubated with probe 1. **B:** *T. brucei* 427 Lister ∆*opb* null mutant parasite genomic DNA incubated with probe 1. **C:** *T. brucei* 427 Lister ∆*opb* null mutant parasite genomic DNA incubated with probe 2. **D:** *T. brucei* 427 Lister ∆*opb* null mutant parasite genomic DNA incubated with probe 3.

4.3.2 ∆*opb* null mutant clones are unable to hydrolyse Z-Arg-Arg-AMC

*Tb*OPB was detected in a western blot with anti-OPB antibodies in wild type *T. b. brucei* Lister 427 parasites, (Fig. 4.8 A, left hand panel, lane 2), and at a lower intensity in the heterozygote (Fig. 4.8 A, left hand panel, lane 3). No *Tb*OPB was detected in the lane loaded with the ∆*opb* null mutant parasites (Fig. 4.8 A, left hand panel, lane 4). By comparison, reprobing the blot with anti-BIP antibodies as a loading control confirmed the presence of equal amounts of BIP in all three samples (Fig. 4.8 A, right hand panel). This confirmed that *TbOPB* gene was deleted in ∆*opb* null mutant parasites. ∆*opb* null mutant parasites were also confirmed in an activity assay using Z-Arg-Arg-AMC (Fig. 4.8, B), the preferred peptide substrate for OPB in *T. b. brucei* (Coetzer *et al.*, 2008). Wild type *T. b. brucei* parasites were able to hydrolyse Z-Arg-Arg-AMC compared to the heterozygote clones ∆*opb::BLA/OPB* and ∆*opb*::*NEO/OPB* which hydrolysed on average half as much of the substrate relative to the wild type parasites (Fig. 4.8 B). The ∆*opb* null mutants hydrolysed less than 5% of the

substrate relative to the wild type (Fig. 4.8 B) that is equivalent to the level of activity obtained with the no lysate control, which may be ascribed to background hydrolysis of the substrate.

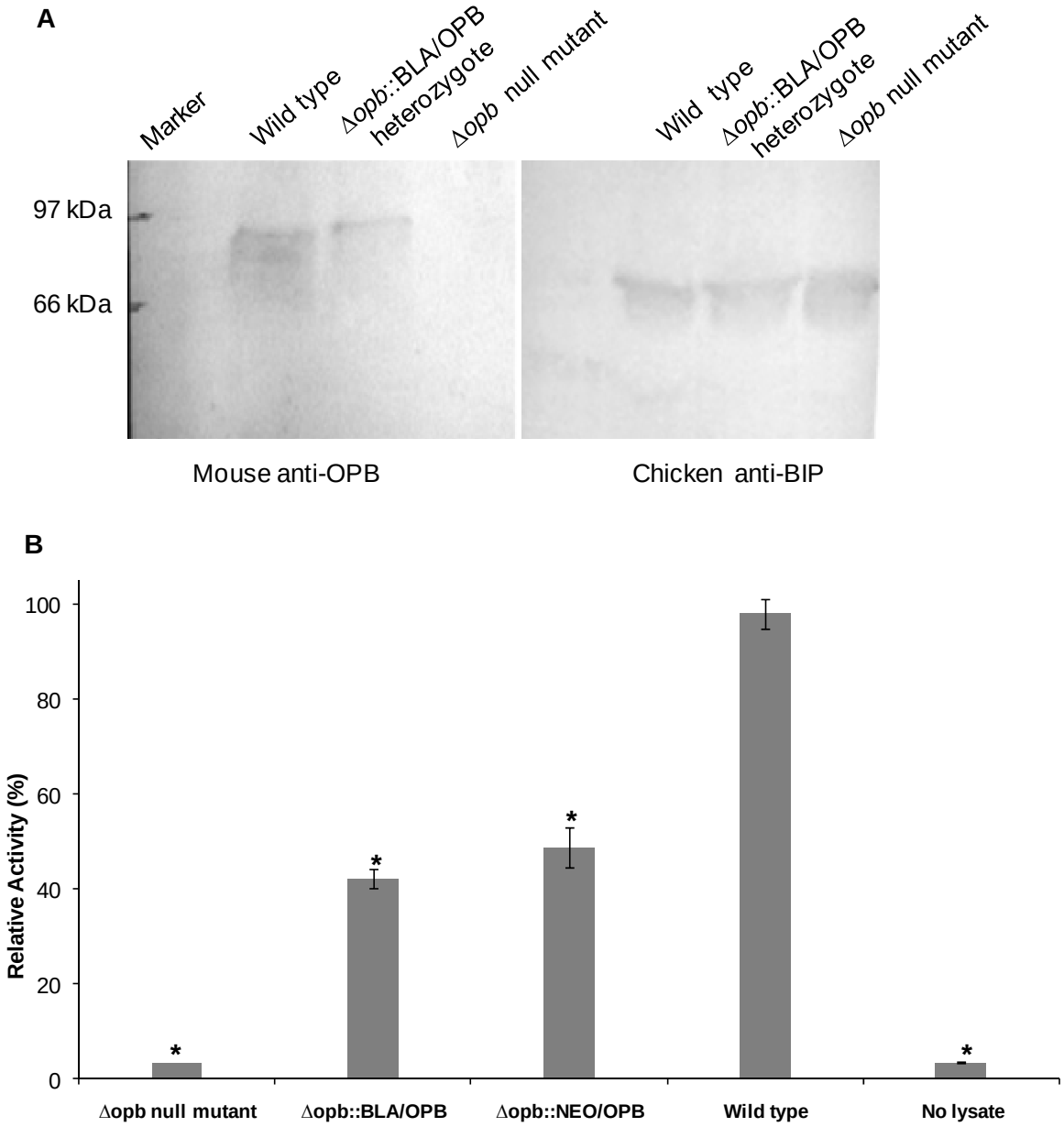
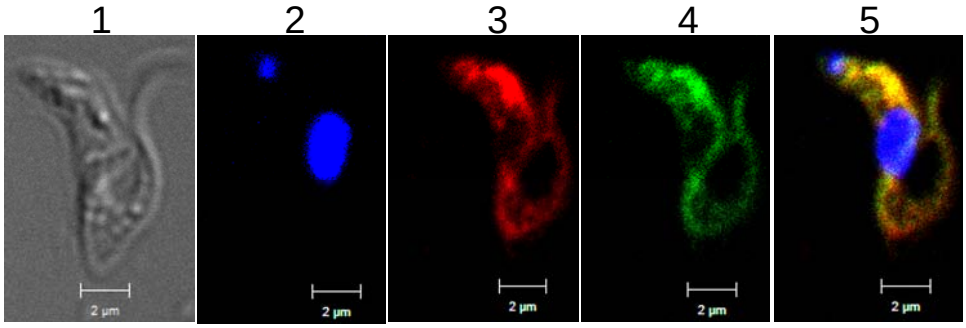


Figure 4.8 Characterisation of Δ*opb* heterozygote and null mutants. A: Mouse anti-OPB antibodies fail to recognise OPB in Δ*opb* null mutants. Polyclonal mouse anti-OPB was used to detect OPB in *T. b. brucei* wild type, heterozygote and Δ*opb* null mutants. Polyclonal chicken anti-BIP antibodies were used as a loading control. Approximately 5 x 10^5 parasites were loaded per well. **B.** Δ*opb* null mutants do not hydrolyse Z-Arg-Arg-AMC. Total parasite lysate (5 µg) from each clone was resuspended in 0.1% (w/v) Brij-35 containing 10 µg/ml soybean trypsin inhibitor, 10 µM E-64 and 1mM EDTA and used to measure the hydrolysis of Z-Arg-Arg-AMC. Fluorescence values obtained relative to those with wild type parasites as a percentage. The mean of triplicate experiments ± SEM was plotted for each experiment. *$p < 0.001$ in comparison to the wild type parasites.

4.3.3 *T. b. brucei* Δ*opb* null mutant parasites do not display any morphological, growth rate or virulence defects

Fluorescence microscopy of wild type and Δ*opb* null mutant parasites showed no differences in morphology where mouse anti-OPB polyclonal antibodies were able to discriminate between null mutants and wild type parasites (Fig. 4.9).

A: Wild type parasites



B: Δ *opb* double knock-out parasites

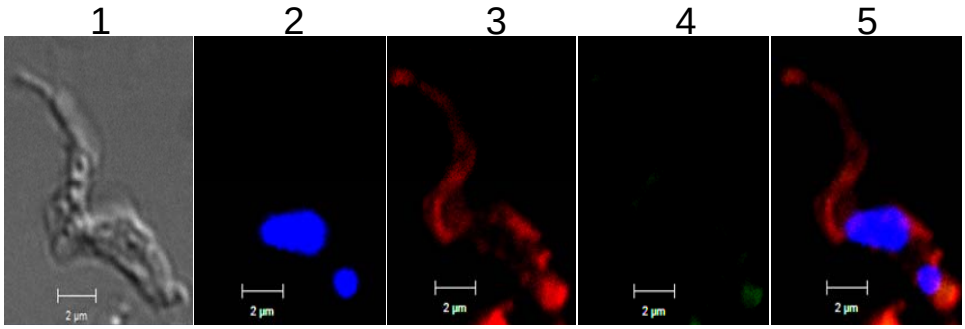


Figure 4.9 *Δopb* null mutant parasites are not recognised by mouse anti-OPB polyclonal antibodies and do not display any morphological differences when compared to wild type *T. b. brucei* Lister 427 strain parasites. Trypanosomes from wild type (**A**) and Δ*opb* null mutant (**B**) cultures were fixed on slides and probed sequentially with polyclonal mouse anti-OPB serum (1:500) and chicken anti-BIP IgY (5 µg/ml) as a control before viewing under a confocal microscope; **panel 1**, DCI light image; **panel 2**, Hoestch nuclear staining; **panel 3**, Cy3 filter for chicken anti-BIP antibodies; **panel 4**, FITC filter for mouse anti-OPB antibodies; **panel 5**, Combined images.

Δ*opb* null mutant clones were assessed for any differences in the rate of division and growth during *in vitro* culture. No differences were seen in the growth of the parasites when compared to the heterozygote or wild type strains (Fig. 4.10 A). The group of mice infected with Δ*opb* null mutant parasites did not survive any longer than the wild type group (Fig. 4.10 C) and parasitaemia counts were not significantly different between both infected groups (Fig. 4.10 B).

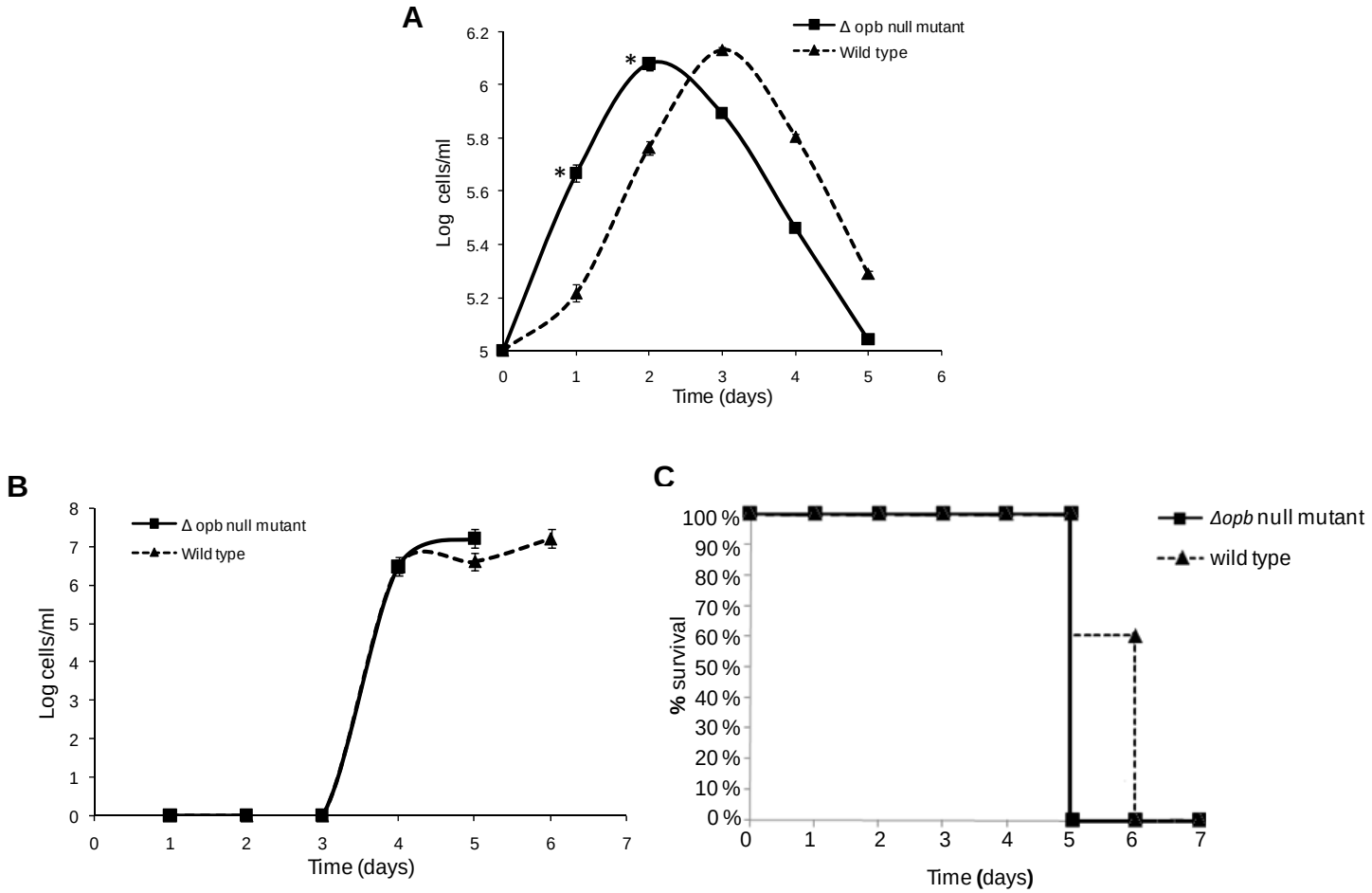


Figure 4.10 Infecting mice with Δ*opb* null mutants had no effect on survival or parasitaemia *in vivo.* **A**: The rate of parasite growth *in vitro* was monitored. *T. brucei* 427 wild type and Δ*opb* null mutant parasites (1 x 10^5/ml of each) were diluted to approximately 1 x 10^5 cells/ml in BSF growth medium and grown at 37°C until saturation (10^6 cells/ml); **p*< 0.001 in comparison to the wild type. **B**: Parasitaemia in mice infected with wild type parasites compared to Δ*opb* null mutant infected mice. Each value is a mean of 5 mice; **p*< 0.001 in comparison to the wild type. **C**: Kaplan-Meier survival analysis for mice infected with *T. b. brucei* wild type parasites compared to Δ*opb* null mutant infected mice (n=5 in each group).

To confirm that *Tb*OPB is not a virulence factor, blood collected from the different infected groups was used to measure residual *Tb*OPB activity *in vivo* (Fig. 4.11). Blood from mice infected with Δ*opb* null mutant parasites was unable to hydrolyze Z-Arg-Arg-AMC when compared to blood from mice infected with the wild type strain (Fig. 4.11). The relative value of fluorescence using blood from Δ*opb* null mutant infected mice was comparable to that from naive mice with no *T. b. brucei* infection.

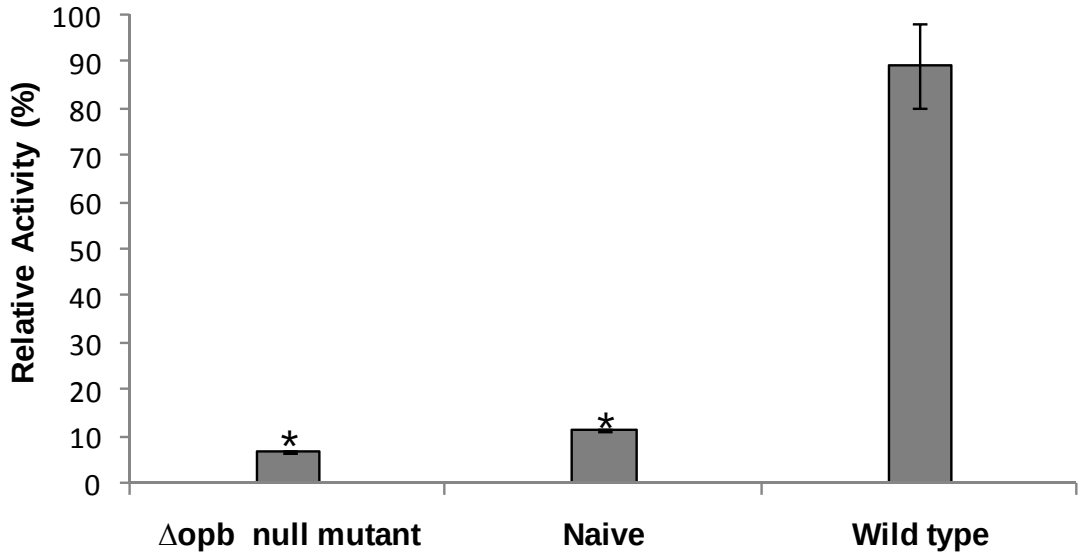


Figure 4.11 Blood from mice infected with Δ*opb* null mutant parasites does not hydrolyse Z-Arg-Arg-AMC. 25 µl of blood from mice infected with wild type parasites, Δ*opb* null mutants and non-infected mice was resuspended in 0.1% (w/v) Brij-35 with 10 µg/ml soybean trypsin inhibitor, 10 µM E-64 and 1mM EDTA and hydrolysis of Z-Arg-Arg-AMC measured. A mean of triplicate fluorescence values were plotted relative to those obtained with the wild type as a percentage for each experiment ± SEM; *p< 0.001 in comparison to the wild type.

T. brucei 427 Lister parasites are known to cross endothelial barriers in mouse testes (Claes *et al.*, 2009). In order to observe if *Tb*OPB plays a role during crossing, immunofluorescence staining of parasites in the testes of mice 5 days after infection with Δ*opb* null mutant parasites showed parasites outside microvascular blood vessels in the interstitial spaces between seminiferous tubules (Fig. 4.12). This confirmed that *T. b. brucei* Δ*opb* null mutant parasites are able to cross endothelial cell barriers in the testes as efficiently as the wild type.

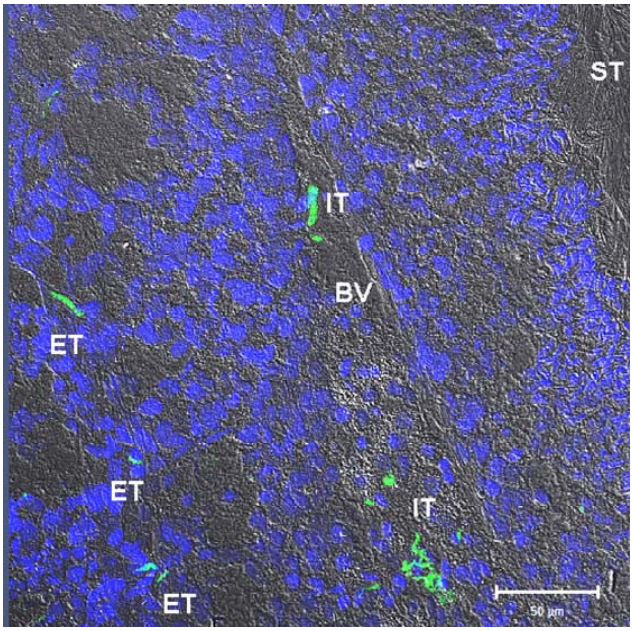


Figure 4.12 Δ*opb* null mutant parasites localise in intravascular and extravascular areas of the mouse testes. Immunofluorescence staining of trypanosomes following incubation with rabbit anti-VSG antibodies in sections of testes from mice 5 days after infection with 1 x 10^6 Δ*opb* null mutant parasites revealed parasites (in green) both inside blood vessels and in interstitial spaces between seminiferous tubules. **IT**, intravascular trypanosomes; **ET**, extravascular trypanosomes; **BV**, blood vessel; **ST**, seminiferous tubule.

4.3.4 *T. b. brucei* Δ*opb* null mutant parasites show significantly elevated levels of prolyl oligopeptidase.

In order to observe how other trypanosomal proteases are affected by knocking out the *TbOPB* gene, proteolytic activity of Δ*opb* null mutant parasite lysates from *in vitro* cultures was assessed on gelatin containing SDS-PAGE gels (Fig. 4.13 A, lane 1) and compared to *T. b. brucei* wild type parasite lysates (Fig. 4.13 A, lane 2). Two areas of gelatin digestion at approximately 200 kDa (Fig. 4.13 A, **i**) and 150 kDa (Fig. 4.13 A, **ii**), were more prominent in the lanes loaded with Δ*opb* null mutant parasite lysate than in lanes containing wild type parasite lysate. In order to identify the class of peptidases showing elevated activity in Δ*opb* null mutants, lysates were separated on duplicate gelatin containing gels and the gels incubated separately with AEBSF (Fig. 4.13 B), pepstatin A (Fig. 4.13 C), E-64 (Fig. 4.13 D), SBTI (Fig. 4.13 E) and TPCK (Fig. 4.13 F). The intensity of bands of hydrolysis on gels incubated with AEBSF, pepstatin A and SBTI were not significantly different from those on the gel incubated in the absence of inhibitors (Fig. 4.13 A). However, in the presence of E-64, digestion of gelatin was only observed at approximately 54 kDa (Fig. 4.13 D). The major

cysteine protease band at approximately 30 kDa was inhibited by both E64 and TPCK (Fig. 4.13 D and F).

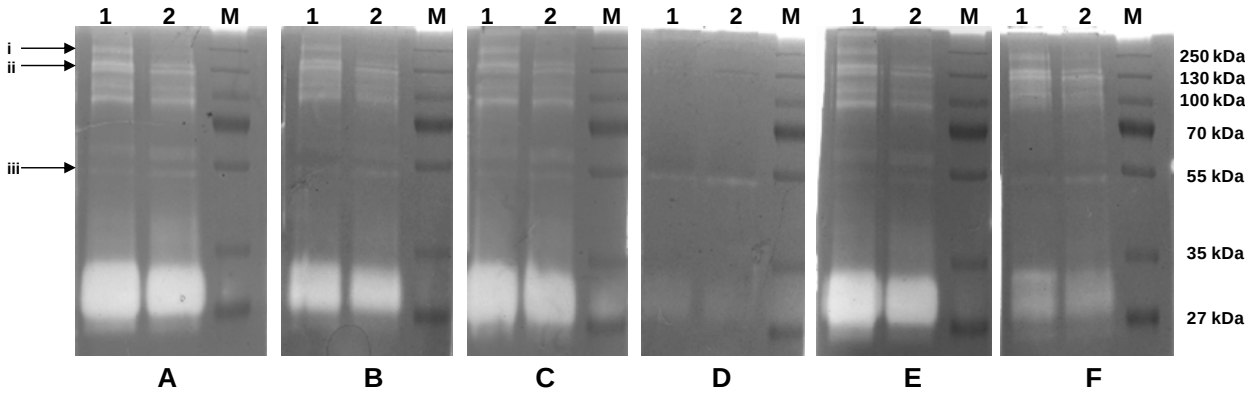


Figure 4.13 *T. b. brucei* Δ*opb* null mutant parasites show elevated levels of cysteine peptidase activity. Lysates from 5 x 10^5 parasites were loaded per well on a gelatin gel and incubated in 0.1 Mes buffer pH 6 overnight with or without inhibitor. **M**: marker; **lane 1,** *T. b. brucei* Δ*opb* null mutant parasite lysate; **lane 2,** *T. b. brucei* wild type parasite lysate. Gels were incubated in 0.1 Mes buffer pH 6 containing **A,** No inhibitor; **B**, 1 mM AEBSF; **C**, 1 µg/ml pepstatin A ; **D**, 1 µM E-64 ; **E**, 100 µg/ml SBTI and **F**, 100 µM TPCK. Gels were stained in freshly prepared 0.1% (m/v) amido black in methanol/acetic acid/water (30:10:60, by vol.) and destained in methanol/ acetic acid/ water (30:10:60, by vol.). Areas at ~200 kDa (**i)** and ~150 kDa (**ii)** at are increased in *T. b. brucei* Δ*opb* null mutant parasites. Only activity at ~54 kDa (**iii)** is not actively inhibited by E-64.

An activity assay using Z-Gly-Pro-AMC and Z-Arg-Gly-Phe-Phe-Pro-4MβNA.HCl was carried out to observe if the levels of prolyl oligopeptidase in *T. brucei* (*Tb*POP) are affected in the absence of *Tb*OPB (Fig. 4.14). Z-Phe-Arg-AMC and Z-Arg-Arg-AMC, the preferred substrates for trypanosomal cysteine proteases and OPB respectively, were also included. Z-Gly-Pro-AMC hydrolysis was increased by more than 500% in Δ*opb* null mutant parasites when compared to the wild type (Fig. 4.14). This markedly elevated activity was confirmed by Z-Arg-Gly-Phe-Phe-Pro-4MβNA.HCl at 200% (Fig. 4.14). Blood from mice infected with Δ*opb* null mutant parasites also displayed increased hydrolysis of Z-Gly-Pro-AMC by 200% when compared to blood from mice infected with wild type parasites (Fig. 4.14). Naive blood from non-infected mice had the highest activity against Z-Gly-Pro-AMC at almost 300% higher than controls due to endogenous prolyl oligopeptidase in the mouse blood (Fig. 4.14). TPCK was more effective in inhibiting *Tb*POP activity in both Δ*opb* null mutant and wild type parasites than AEBSF (Fig. 4.14). A slightly elevated level at approximately 10% of Z-Phe-Arg-AMC hydrolysis was observed in Δ*opb* null mutant parasites when compared to the wild type confirming the increased cysteine protease activity observed in the zymograms (Fig. 4.14). E-64 inhibited cysteine protease activity in both Δ*opb* null mutant and wild type parasites (Fig. 4.14).

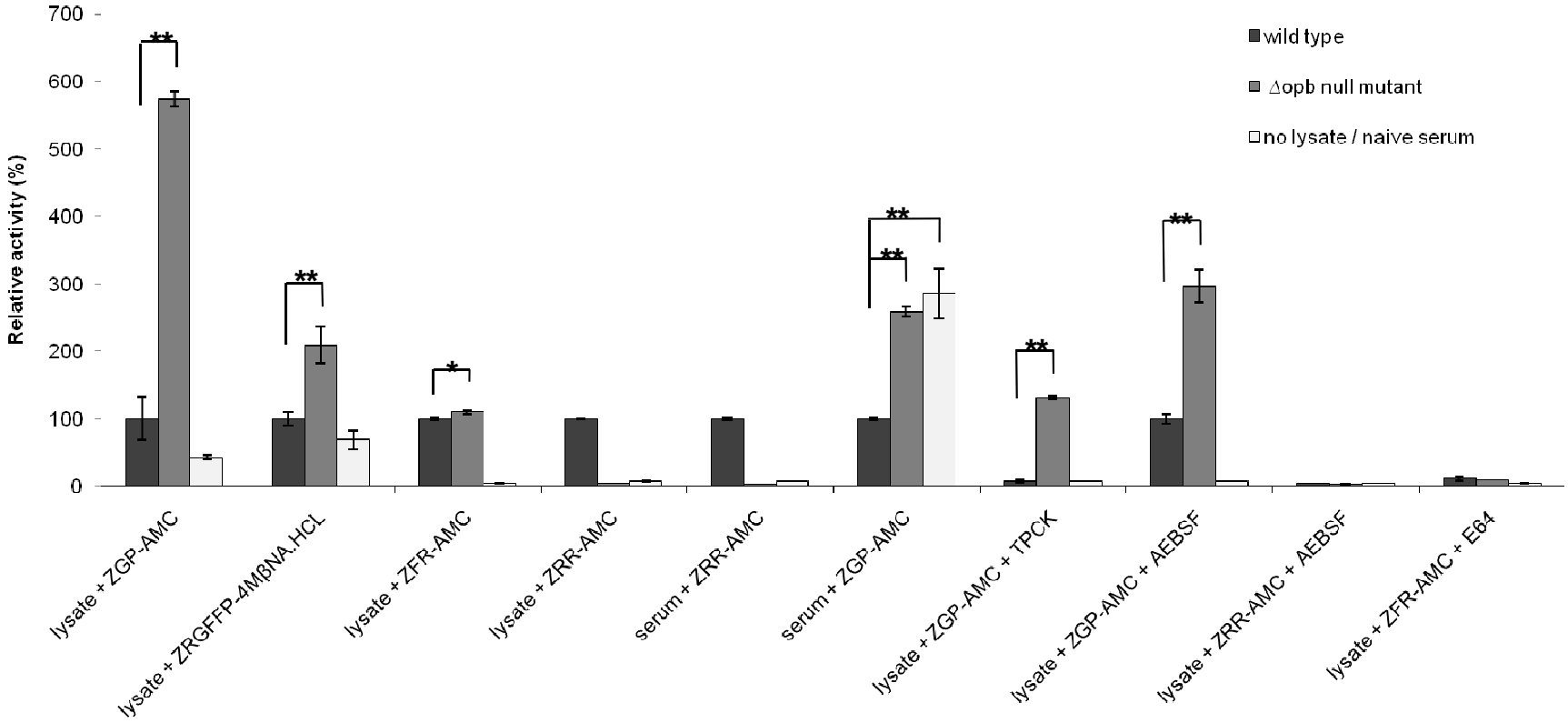


Figure 4.14 Δ*opb* null mutants hydrolyse significantly more Z-Gly-Pro-AMC than wild type parasites. Total parasite lysate from each clone (5 µg) or 25 µl of blood from infected mice was resuspended in 0.1% (w/v) Brij-35 with 10 µg/ml SBTI, 1 µg/ml of pepstatin A, and 1 mM EDTA and used to measure hydrolysis of either Z-Arg-Arg-AMC specific for *Tb*OPB, Z-Phe-Arg-AMC specific for cysteine peptidases and Z-Gly-Pro-AMC or Z-Arg-Gly-Phe-Phe-Pro-4MβNA.HCL specific for *Tb*POP as peptide substrates. A mean of triplicate fluorescence values was plotted relative to those obtained with the wild type parasites as a percentage for each experiment ± SEM. *p< 0.05 and **p<0.001 in comparison to the wild type parasites. Identical triplicates were incubated with the inhibitors 10 µM E-64, 100 µM TPCK and 1 mM AEBSF to confirm that the increase in activity when using Z-Gly-Pro-AMC as a substrate is attributed to prolyl oligopeptidase in Δ*opb* null mutant parasites.

To assess if this increase in hydrolysis of substrates containing a prolyl residue in P_1 was due to an increased level of *Tb*POP expression in Δ*opb* null mutant parasites, a western blot using anti-POP mouse serum revealed that *Tb*POP levels remained unchanged in Δ*opb* null mutant parasites compared to wild type parasites (Fig. 4.15, left hand panel). Reprobing the blot with anti-BIP antibodies as a loading control confirmed the presence of equal amounts of BIP in both parasite lysates (Fig. 4.15, right hand panel).

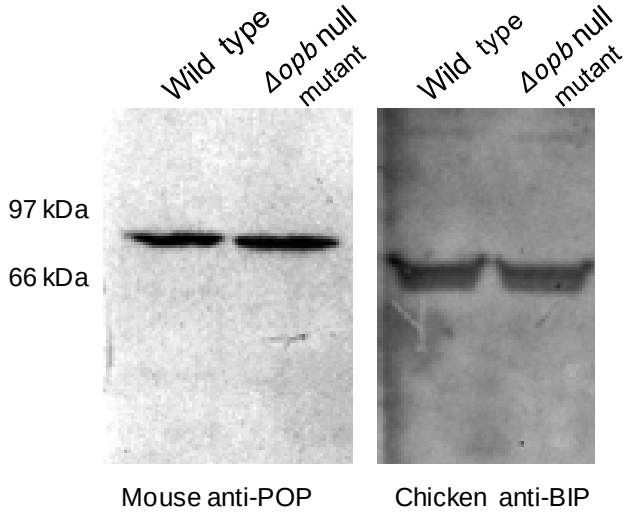


Figure 4.15 Measuring protein levels of POP in *T. brucei* 427 wild type parasites and Δ*opb* null mutants. Mouse anti-POP antibodies recognise POP in Δ*opb* null mutants at similar levels to POP in *T. brucei* wild type parasites. Polyclonal mouse anti-POP serum was used at a dilution of 1:300. Polyclonal chicken anti-BIP antibodies were used as a loading control. Approximately 5 x 10^5 parasites were loaded per well.

4.4 DISCUSSION

Since *Tb*OPB is expressed as an active cytosolic peptidase at all stages of the *T. brucei* life cycle (Mbawa *et al.*, 1991; Morty *et al.*, 1999b), it is likely to have more than one function given the varied environments the parasite endures in different hosts. Obtaining Δ*opb* null mutants in this experiment confirms that *Tb*OPB does not play an essential role in parasite survival as has been previously observed in other trypanosomatid parasites such as *T. cruzi* (Caler *et al.*, 1998) and *Leishmania spp.* (Munday *et al.*, 2011; Swenerton *et al.*, 2011). This finding is confirmed by similar growth kinetics in Δ*opb* null mutants when compared to those of the wild type strain even though Δ*opb* null mutants showed no activity against Z-Arg-Arg-AMC. Unfortunately, it was only possible to obtain OPB heterozygotes in *T. congolense* IL3000. Previous work has shown that knocking out OPB has a minimal effect on the establishment of disease in *T. cruzi* and *Leishmania* (Caler *et al.,* 1998; Munday *et al.,* 2011;

Swenerton *et al.,* 2011) and this was also observed in the present study with *T. brucei* ∆*opb* null mutants that were as virulent and pathogenic as the wild type parasites. This result correlates well with observations in studies with intracellular trypanosomatids where strong phenotypes in the absence of OPB were mostly associated with host cell invasion rather than with the establishment of disease in mammalian hosts.

Because *T. brucei* is an exclusively extracellular parasite it is not surprising that a specific phenotype was not observed in infected mice without active *Tb*OPB in the bloodstream. *T. brucei* is an exclusively extracellular parasite that releases OPB after lysis in the host (Morty *et al.*, 2001; Coetzer *et al.*, 2008). A mouse infection study was carried out in order to assess the possible role that OPB plays in an infection involving extracellular kinetoplastids. There were no significant differences in parasitaemia or survival between mice infected with either *T. brucei* Δ*opb* null mutants or wild type parasites. It must be emphasised that although mice have been extensively used as a model for trypanosome infections (Antoine-Moussiaux *et al.*, 2008) and (Magez and Caljon, 2011), extrapolations to the natural infection in cattle should be done with caution. *T. b. brucei* 427 Lister strain infections in mice are used as a model for the acute disease in ruminants (Claes *et al.*, 2009), with death occurring within 6–8 days after an intraperitoneal injection of 1×10^3 parasites. It would be difficult to evaluate the role of *Tb*OPB in mice infected in this manner when compared to a natural infection, where the number of parasites infecting the host would be less when delivered by a tsetse bite and the resulting infection is chronic (Magez *et al.*, 2010). The *in vivo* functions of *Tb*OPB can only be effectively studied in a ruminant host where a more natural chronic infection is possible, and thus, it would be inappropriate to rule out *Tb*OPB as a virulence and pathogenic factor using the results observed in this study.

T. brucei parasites are known to cross blood barriers in immuno privileged organs such as the brain (Grab *et al.*, 2004; Grab and Kennedy, 2008; Grab *et al.*, 2009) and testes (Claes *et al.*, 2009), with cysteine proteases involved in facilitating this process (Nikolskaia *et al.*, 2006a; Abdulla *et al.*, 2008). It has been previously suggested that *Tb*OPB may have a role to play in crossing endothelial barriers in microvascular vessels (Lonsdale-Eccles and Grab, 2002), and to investigate this thin sections of ∆*opb T. brucei* infected mouse testes were probed with rabbit anti-VSG antibodies. Testes were chosen for sectioning because *T. b. brucei* Lister 427 localises and displays enhanced parasite development in this organ (Claes *et al.*, 2009). Parasites were identified outside blood vessels in the interstitial spaces between

seminiferous tubules. This confirmed that *Tb*OPB is not involved in assisting parasites in crossing endothelial cell blood barriers into tissue.

With no definite phenotype identified for *Δopb T. brucei* parasites, an investigation of how the activity of other *T. b. brucei* peptidases is affected by the absence of *Tb*OPB was undertaken. This would indicate if there are any overlapping functions between *Tb*OPB and other proteases within the same serine peptidase family and also across peptidase families. Cathepsin L-cysteine peptidases in *F. hepatica* belong to a large family divided into five groups with different substrate specificities that overlap in function (Robinson *et al.*, 2008). The expression of these different peptidase classes is tightly regulated according to the life cycle stage and environment where *F. hepatica* is found. Other instances of overlapping functions across different peptidase families include plasminogen and different classes of matrix metalloproteinases (MMPs) (Lund *et al.*, 1999; Solberg *et al.*, 2003). Using gelatin zymograms in the presence and absence of class-specific inhibitors, we were able to identify an increase in the levels of cysteine peptidase activity in *Δopb T. b. brucei* parasites when compared to the wild type parasites.

It has been suggested for *Leishmania*, that up-regulation of other *OPB*-like genes might compensate for the loss of OPB (de Matos Guedes *et al.*, 2008; Munday *et al.*, 2011). Because oligopeptidases do not hydrolyse gelatin, peptidolytic activity assays with a proline residue in the P_1 position (Bastos *et al.*, 2010) was used to assess activity levels of *Tb*POP, a well characterised prolyl oligopeptidase that belongs to the same peptidase family as OPB and potentially hydrolyse the same physiological peptide hormone substrates that contain both Arg/Lys and Pro residues (Table 4.2). Unlike other serine oligopeptidases, *Tb*POP is poorly inhibited by AEBSF and more effectively inhibited by TPCK (Bastos *et al.*, 2010). These unique characteristics of the peptidase were evident in assays using Arg-Gly-Pro-AMC where the levels of POP-activity were increased by more than 500% in *Δopb* null mutants compared to wild type parasites and the activity was more effectively inhibited by TPCK than by AEBSF. Z-Arg-Gly-Phe-Phe-Pro-4MβNA·HCl was not as effective as the AMC substrate because a higher background value was observed in the no lysate control. An accompanying increase in *Tb*POP protein levels in Δ*opb* null mutants was not observed when compared to *T. brucei* wild type lysates, signifying that either another POP-like peptidase was responsible for the observed increase in activity, or that *Tb*OPB is able to activate an inhibitor specific for *Tb*POP, thus leading to an increase in *Tb*POP activity in its absence.

Table 4.2 Hydrolysis of peptide hormones by *Tb*OPB and *Tb*POP

Peptide	*Tb*OPB hydrolysis sites	*Tb*POP hydrolysis sites
ANF	SLR↓ R↓ SSCFGGR↓ IDRIGAQSGLGCNSFR↓ Y [a]	SLRRSSCFGGRIDRIGA↓ QSGLGCNSFRY
β-endorphin	YGGFMTSEK↓ SQTPLVTLFK↓ NAIIK↓ NAYK↓ K↓ GE	YGGFMTSEKSQTP↓ LVTLFKNA↓ IIKNAYKKGE [b]
Neurotensin	pELYENK↓ PR↓ R↓ PYIL-NH_2	pELYENKP↓ RRP↓ YIL-NH_2 [b]
GnRH	pEHWSYGLR↓ PG-NH_2 [c]	pEHWSYGLRP↓ G-NH_2 [b]
Bradykinin	RPPGFSPFR↓ -NH_2	RP↓ P↓ GFSP↓ FR-NH_2 [b]
TRH	pEHP-NH_2	pEHP↓ -NH_2 [b]

[a]Morty *et al.* (2005a); [b]Bastos *et al.* (2010); [c]Tetaert *et al.* (1993).

Previous experiments involving the inhibition of *Tb*OPB by trypanocidal drugs (Morty *et al.*, 1998) and phosphonate diphenyl esters (Morty *et al.*, 2000) have shown it to be a major target of these compounds. Active site labelling identified a band at 80 kDa using fluorescent irreversible inhibitors with an arginine residue or arginine analogue in P_1 (Morty *et al.*, 2000) although native *Tb*POP is also approximately 80 kDa in size (Bastos *et al.*, 2010), it would not have been labelled by the fluorescent inhibitor with this P1 specificity. However, as shown in the same study (Morty *et al.*, 2000), these fluorescent irreversible inhibitors also targeted other peptidases in *T. b. brucei* lysates: the peptidyl phosphonate diphenyl ester targeted a minor 45 kDa protein, while 120 and 68 kDa proteins were targeted by biotin-Arg-CH_2Cl. By acting on multiple targets in trypanosomes these inhibitors would therefore be able to kill trypanosomes in culture and clear *T. brucei* parasites in infected mice (Morty *et al.*, 2000), whereas specific deletion of this *TbOPB* gene appears not to be sufficient to affect the viability of the Δ*opb* null mutant trypanosomes. Several potential peptide hormone substrates all have alanine or proline and arginine or lysine residues that can be cleaved by *Tb*POP and *Tb*OPB, respectively (Table 4.2). This would imply that in the absence of *Tb*OPB, *Tb*POP or a closely related POP-like peptidase would still be capable of hydrolysing these hormones and other substrates at a higher turnover.

In conclusion, *T. brucei Δopb* null mutant parasites were generated with no OPB activity, but their *Tb*POP activity was increased significantly. Since *T. brucei* serine oligopeptidases have the ability to hydrolyse the same peptide hormones, in the absence of one peptidase (*Tb*OPB), other peptidases (such as *Tb*POP) can be up-regulated in activity to take over that function. When controlling trypanosomosis, it will be necessary to consider more than one target for drug and vaccine development in order to give full protection to the disease.

CHAPTER 5

GENERAL DISCUSSION

African animal trypanosomosis is a disease in cattle that poses a major impediment to agricultural development in sub-Saharan Africa and costs the continent USD $ 1 - 5 billion annually in economic losses (Kristjanson *et al.*, 1999; Baral, 2010). Chemotherapy, the main mode of treatment available to subsistence farmers in Africa (McDermott and Coleman, 2001), is fast becoming a non-viable option due to the development of resistance in trypanosomes to commonly used drugs in the region (Delespaux and de Koning, 2007). The development of a vaccine as an alternative to chemotherapy has been hampered by antigenic variation in trypanosomes, which makes it difficult to use a VSG based vaccine (McCulloch and Horn, 2009). An alternative that has been suggested is to target pathogenic products released during an infection in a strategy referred to as an 'anti-disease' approach (Authié *et al.*, 1993b; Schofield, 2007; Antoine-Moussiaux *et al.*, 2009). In trypanosomes, several toxins that cause harm are released into the host circulatory system. These include lipids, hemolysins, mitogens, complement activating and inflammatory factors, hepatotoxins and peptidases (Tizard *et al.*, 1978; Antoine-Moussiaux *et al.*, 2009). Phospholipase A_1 from *T. cruzi* has been shown to mediate the breakdown of host phosphatidylcholine resulting in an increase of free fatty acids levels in circulation that have the potential to disrupt membrane phospholipid bilayers (Wainszelbaum *et al.*, 2001). Sialidases released during trypanosome infection are responsible for pathogenesis by altering red blood cell surfaces resulting in phagocytosis and subsequent anaemia (Nok and Balogun, 2003). Other pathological factors released during trypanosome infection include proline racemase that acts as a B cell mitogen leading to polyclonal B cell activation thus diverting parasite specific responses in *T. cruzi* infections (Reina-San-Martin *et al.*, 2000). The abnormal activation of host complement is caused by trypanosome products such as GP63 surface protease and calreticulin leading to tissue degradation and inflammation (Zambrano-Villa *et al.*, 2002; Jameel M, 2004; Ferreira *et al.*, 2005). Peptidases that have been implicated in pathogenesis include cathepsin L-like peptidases (CATLs) and oligopeptidases (Mottram *et al.*, 1998; Authié *et al.*, 2001; Coetzer *et al.*, 2008).

Trypanotolerant cattle, such as the N'dama longhorn, are able to remain productive under natural trypanosome challenge by controlling parasitaemia and anaemia, and by producing antibodies specific to pathogenic invariant trypanosome antigens (Naessens *et al.*, 2002).

Although these antigen-specific antibodies are not responsible for controlling parasite numbers during infection, they may be important for neutralising the pathogenic products released by infecting trypanosomes. The two peptidases studied in this work, *Tco*CATL (congopain) and OPB (oligopeptidase B), have been implicated in causing pathogenesis during natural trypanosome infections. A high IgG response to *Tco*CATL has been observed in trypanotolerant cattle as opposed to trypanosusceptible cattle (Authié *et al.*, 1992; Authié *et al.*, 1993a; Authié *et al.*, 1993b), indicating that an immune response against the peptidase is necessary to control symptoms associated with the disease. Rats, experimentally infected with *T. brucei*, have active OPB released into their bloodstream by dead and dying parasites which is responsible for catalysing the hydrolysis of peptide hormones leading to homeostatic associated pathologies such as hypervolemia and circulatory lesions (Morty *et al.*, 2001; Coetzer *et al.*, 2008). The main aim of this study was to investigate the function of two trypanosome peptidases that have been identified as pathogenic factors during infection, namely *Tco*CATL and oligopeptidase B in *T. brucei* and *T. congolense*, using gene disruption.

*Tco*CATL is a member of the cathepsin L-like subfamily of the C1 family within clan CA of cysteine peptidases (Rawlings *et al.*, 1994; Rawlings *et al.*, 2008). Members of this subfamily have the 'ERFNIN' conserved motif in their pro-region, substrate specificity defined by the S_2 pocket and like other members of clan CA, are sensitive to inhibition by the cysteine peptidase inhibitor E64 (Caffrey and Steverding, 2009). *Tco*CATL and *Tcr*CATL have broad pH profiles that include physiological pH (Mbawa *et al.*, 1992; Chagas *et al.*, 1997; Pillay *et al.*, 2010). The ability of parasite CATL-like peptidases to remain active at neutral pH enables them to carry out various extra-lysosomal functions necessary for survival in the mammalian host. As described in Section 2.3.3 of this work, inducing RNAi against *Tco*CATL *in vivo* led to a reduction of circulating parasites in mice. It was, however, not possible to induce *Tco*CATL RNAi in procyclics, the only *T. congolense* life cycle stage that was successfully grown *in vitro* during the course of these experiments. Considering that *Tco*CATL is only expressed maximally during the bloodstream stage of the *T. congolense* life cycle (Mbawa *et al.*, 1991; Mbawa *et al.*, 1992; Pillay, 2011), targeting *Tco*CATL at the procyclic stage, would not have provided any functional information. An attempt to grow the bloodstream form of the parasite *in vitro* was not successful, and modifications such as using fresh goat serum and red blood cell lysate were made following a recently described protocol (Coustou *et al.*, 2010). As an alternative to *in vitro* studies, it was decided to use transformed procyclic parasites that had been differentiated into metacyclics to infect mice and induce RNAi *in vivo*,

similar to a study carried out with *Tb*CATL (Abdulla *et al.*, 2008). Groups of mice infected with parasites induced for *Tco*CATL RNAi displayed a lower parasite burden for an extended period of time that could be correlated to reduced *Tco*CATL activity in blood. This phenotype can be attributed to a defect in different mechanisms where CATL peptidases are crucial for parasite survival such as nutrition, immunoevasion, virulence and the processing and activation of host and parasite proteins (Sajid and McKerrow, 2002). A summary of these CATL functions in different kinetoplastid parasites is shown in Table 5.1.

CATL-like peptidases are responsible for the degradation of host proteins that provide nutrition in different parasites (Sajid and McKerrow, 2002). In *P. falciparum*, falcipain-2, a CATL-like cysteine peptidase, is responsible for degrading host haemoglobin for nutrition in early to mid stage trophozoites (Sijwali and Rosenthal, 2004). This function was described after knocking out the gene transcribing falcipain-2, and an accumulation was observed of undigested haemoglobin in the parasite food vacuole (Sijwali and Rosenthal, 2004). A similar phenotype was also observed in parasites treated with E64 which also resulted in the accumulation of undigested haemoglobin (Gamboa de Dominguez and Rosenthal, 1996). However, in falcipain-2 null mutants, the phenotype is transient and late trophozoites recovered as they developed into the schizont stage. This has been attributed to the presence of a second gene coding for falcipain-2 that is transcribed in late stage trophozoites (Sijwali and Rosenthal, 2004). The characterisation of recombinant CATL from *S. mansoni* (*Sm*CL1) revealed that the peptidase is also capable of degrading haemoglobin (Brady *et al.*, 1999). Interestingly, *Sm*CL1 is only able to degrade haemoglobin at pH 4-4.5, indicating that it is most active in the schistosome gut where ingested haemoglobin is abundant (Brady *et al.*, 1999). In *T. brucei*, the degradation of transferrin has been attributed to *Tb*CATB rather than *Tb*CATL (Mackey *et al.*, 2004). Native *Tb*CATL isolated from infected mice hydrolyses fibrinogen, BSA and rabbit IgG at physiological pH *in vitro* (Troeberg *et al.*, 1996b). However, it has also been noted that *Tb*CATL catalytic activity *in vivo* is inhibited by host cystatins (Nwagwu *et al.*, 1988; Troeberg *et al.*, 1996b). *T. brucei* parasites targeted for CATB knock-down using RNAi showed an accumulation of FITC-transferrin in RNAi induced parasites (Mackey *et al.*, 2004). Bloodstream form parasites acquire iron from the host, and the reduced levels of *Tb*CATB led to parasite death *in vitro* and during infection (Mackey *et al.*, 2004; Abdulla *et al.*, 2008). *Tco*CATL may also be responsible for the degradation of host proteins due to its *in vitro* activity on proteins such as BSA, fibrinogen and VSG (Mbawa *et al.*, 1992). The role that *Tco*CATL plays in degrading host proteins may be in conjunction with

other peptidases where more than one peptidase can be involved in processing as observed in *P. falciparum* parasites. The degradation of haemoglobin in malaria parasites requires plasmepsins that hydrolyse haemoglobin to globin polypeptides (Silva *et al.*, 1996), falcipain-2 and 3 which hydrolyse the polypeptides to 10-20 residue-long oligopeptides (Sijwali *et al.*, 2001; Sijwali and Rosenthal, 2004), falcilysin that further catalyse the hydrolysis of oligopeptides to 5-10 residue long peptides (Murata and Goldberg, 2003) and finally dipeptidyl aminopeptidase 1 that cleaves the short peptides to dipeptides for amino acid uptake (Klemba *et al.*, 2004).

Trypanosome CATLs are also involved in protein processing and activation (Caffrey *et al.*, 2011). Taking into account that *T. congolense* is an extracellular parasite that is strictly confined to the host blood circulation (Ojok *et al.*, 2002), protein processing becomes important in relation to evading the host immune response to which the parasite is constantly exposed. One method through which trypanosomes are able to escape immune targeting by the host is by internalising and degrading trypanosome specific antibodies that are bound to surface VSG, as has been observed in *T. cruzi* (Teixeira and Santana, 1989). In *T. brucei*, deletion of a gene that codes for the endogenous cysteine peptidase inhibitor (ICP) resulted in parasites with a three-fold increase in *Tb*CATL activity, and the degradation of VSG-IgG complexes was more efficient when compared to the wild type (Santos *et al.*, 2007). Furthermore, the increase in VSG-bound IgG degradation in parasites without ICP was specifically inhibited when K11777 and E-64 (both inhibitors of *Tb*CATL) were added to parasites labelled with anti-VSG 221 antibodies *in vitro* (Santos *et al.*, 2007). These observations suggest that the increased degradation of VSG-IgG complexes by the parasites lacking ICP was due to higher *Tb*CATL activity in the absence of ICP, thus implicating *Tb*CATL in the breakdown of VSG-antibody complexes in the lysosome. In *T. cruzi*, *Tcr*CATL hydrolyses high molecular weight kininogen into kinins and this process is enhanced in the presence of heparin sulfate (Lima *et al.*, 2002). Kinins released through the action of *Tcr*CATL activate local endothelial or smooth muscle cells through kinin receptors, leading to inflammatory responses aiding parasite cell invasion and causing symptoms associated with heart disease, a hallmark of chagas disease (Scharfstein *et al.*, 2000). The unusual processing of the merozoite surface protein-1 (MSP-1) by falcipain seems to occur in *P. falciparum* to facilitate merozoite release from the erythrocyte (Sajid and McKerrow, 2002). Falstatin, an endogenous inhibitor for falcipain, regulates falcipain-2 activity in late-stage merozoites to modulate the hydrolysis of MSP-1 (Pandey *et al.*, 2006).

Table 5.1. Summary of kinetoplastid clan CA family C1 cysteine peptidases [Adapted from Caffrey *et al.* (2011)].

Species	Peptidase Name(s)	Life stage(s) expressed	Sub-cellular distribution	Assigned Functions	References
T. congolense	*Tco*CATL1 (CP1)	Bloodstream	Flagellar pocket, Lysosome	Trypanotolerance	i, ii
	*Tco*CATL2 (CP2, congopain, trypanopain-Tc)	Bloodstream	Lysosome	Trypanotolerance, degradation of endocytosed proteins	ii, iii, iv, v
	*Tco*CATB1-13	Bloodstream (*Tco*CATB1-5,12) Procyclics (*Tco*CATB1,2,6,11,12)	Lysosome	Degradation of endocytosed proteins	vi, vii
T. brucei	*Tb*CATL (brucipain, trypanopain-Tb)	Bloodstream	Lysosome	Degradation of anti-VSG IgG and host proteins, crossing blood-brain barrier, disease progression in mice, degradation of host proteins	vi, viii, ix, x, xi, xii
	*Tb*CATB (tbcatB)	All life stages	Lysosome	Degradation of transferrin, degradation of endocytosed proteins, disease in mice	xiii, xiv
Leishmania spp.	*Lm*CATL-A	All life stages	Lysosome	Macrophage infection, virulence, potentiates Th-2 response in host	xv, xvi
	*Lm*CATL-B1-2	Promastigotes, metacyclics	Micro-vesicular tubules, flagellar pocket, lysosome	Macrophage infection, virulence, lesion progression, host immune modulator (switches host immune response from Th-1 to Th-2)	xvii, xviii, xix, xx, xxi
	*Lm*CATL-B3-18	Amastigotes	Lysosome	Macrophage infection, virulence	xxii, xxiii
T. cruzi	*Tcr*CATL1 (cruzipain)	Epimastigotes, trypomastigotes, amastigotes	Golgi, reservosomes, flagellar pocket, vesicles	Generation of kinins, inflammation, host cell signalling, host cell invasion.	xxiv, xxv, xxvi
	*Tcr*CATL2 (cruzipain 1)	Trypomastigotes	Undetermined	Host cell invasion	xxvii

[i]Fish *et al.* (1995), [ii]Boulangé *et al.* (2001), [iii]Mbawa *et al.* (1992), [iv]Authié (1994), [v]Chagas *et al.* (1997), [vi]Mbawa *et al.* (1991), [vii]Mendoza-Palomares *et al.* (2008), [viii]Troeberg *et al.* (1996b), [ix]Nikolskaia *et al.* (2006a), [x]Santos *et al.* (2007), [xi]Abdulla *et al.* (2008), [xii]Grab *et al.* (2009), [xiii]Tazeh *et al.* (2009), [xiv]Mackey *et al.* (2004), [xv]O'Brien *et al.* (2008), [xvi]Zhang *et al.* (2001), [xvii]Onishi *et al.* (2004), [xviii]Mottram *et al.* (1996), [xix]Alexander *et al.* (1998), [xx]Mottram *et al.* (1998), [xxi]Pollock *et al.* (2003), [xxii]Mottram *et al.* (2004), [xxiii]Saravia *et al.* (2006), [xxiv]Scharfstein *et al.* (2000), [xxv]Lima *et al.* (2001), [xxvi]Lima *et al.* (2002), [xxvii]Teixeira and Santana (1989).

Although a calcium dependent serine peptidase has been implicated in the processing of MSP-1 during erythrocyte rupture and invasion (Harris *et al.*, 2005), inhibition of falstatin by specific antibodies at the late schizont stage paradoxically blocked erythrocyte invasion suggesting that falcipain hydrolyses peptidases necessary for cell invasion (Pandey *et al.*, 2006).

Kinetoplastid parasites are also able to evade the immune system by modulating the host immune system through the action of CATLs (Mottram *et al.*, 1998). In wild type *L. mexicana infections,* parasites modulate the host to produce an IL-4 mediated Th-2 response as opposed to an IL-12 mediated Th-1 response (Alexander *et al.*, 1998; Mottram *et al.*, 1998; Saravia *et al.*, 2006). CATLs in *L. mexicana* are encoded for by two different gene loci; *LmcCATL-A* (*cpa*), a single-copy gene encoding a cathepsin L-like cysteine peptidase expressed at lower levels (Mottram *et al.*, 1992); and *LmcCATL-B* (*cpb*), a multi-copy gene that codes for the major CATL-like cysteine peptidase expressed in high levels during the intracellular amastigote stage of the parasite life cycle (Souza *et al.*, 1992). Deleting the *LmcCATL-A* and *LmcCATL-B* loci that code for both forms of *Leishmania* cysteine peptidases resulted in a shift of the host immune response from a predominantly Th-2 response in mice infected with wild type parasites, to a Th-1 response in mice infected with mutant parasites lacking both *Lmc*CATL genes (Alexander *et al.*, 1998). An IL-12 mediated Th-1 response in mice infected with CATL mutant *Leishmania* produces IFN-γ, that in turn induces macrophages to express nitric oxide which is lethal to infecting parasites (Heinzel *et al.*, 1993). Hamsters infected with *L. mexicana* parasites deficient in both forms of *Lmc*CATLs displayed a lower parasite burden, had smaller foot pad lesions and a slower progression of the disease (Saravia *et al.*, 2006). Interestingly, *Lmc*CATL-B seems to contribute more to immunoevasion than *Lmc*CATL-A. Immunising mice with *Lmc*CATL-B2.8, a product of the *Lmc*CATL-B multi-copy gene locus, induced a strong IL-4 mediated Th-2 response, with increased IL-4 and IgE production (Pollock *et al.*, 2003). The contribution of *Lmc*CATL-A cannot however be discounted as *Lmc*CATL-A/B mutants consistently display lower infectivity *in vivo* and *in vitro* and induce stronger IL-4 cytokine production than *Lmc*CATL-B mutants (Saravia *et al.*, 2006).

The reduction of parasites induced for *Tco*CATL RNAi *in vivo* described in Section 2.3.3 of this thesis could also be attributed to a loss of virulence, a role that has been described for kinetoplastid CATLs [reviewed in Caffrey *et al.* (2011)]. In *Leishmania mexicana*,

promastigote mutant parasites without the *LmcCATL-B* gene locus are less able to infect macrophages *in vitro* than wild-type parasites (Frame *et al.*, 2000). In contrast, amastigote mutant *LmcCATL-B* parasites were able to survive in macrophages *in vitro* pointing to a stage-specific difference (Frame *et al.,* 2000). Both life stage mutants were, however, able to cause lesions in BALB/c mice but at a much slower rate than wild type parasites. Various roles for CATL in trypanosome virulence were also observed in *T. brucei*. *Tb*CATL is necessary for the parasite to traverse the blood-brain barrier (Nikolskaia *et al.*, 2006a; Abdulla *et al.*, 2008; Grab and Kennedy, 2008). This is a calcium dependent process that is inhibited in an *in vitro* model by K11777, a specific CATL inhibitor (Nikolskaia *et al.*, 2006a). An *in vivo* model that induces *Tb*CATL RNAi, similar to the methods used in the present work, delayed the onset of disease with 50% of the mice used in the study surviving for an extra 60 days when compared to the wild type infection (Abdulla *et al.*, 2008). Parasites induced for *Tb*CATL RNAi *in vitro* were also significantly slower at crossing a blood-brain barrier model (Abdulla *et al.*, 2008). Taken together, these results suggest that *Tb*CATL is assisting the parasite in crossing the blood brain barrier via a calcium dependent process (Grab and Kennedy, 2008). Inhibitors targeting *Tb*CATL *in vivo* would be important when formulating drugs that can be applicable in late stage cases of the disease when the parasite crosses into the brain (Grab and Kennedy, 2008; Grab *et al.*, 2009).

Oligopeptidase B in *T. brucei* (*Tb*OPB) belongs to the S9 prolyl oligopeptidase family, which is grouped in clan SC of serine peptidases in the MEROPS database (Barrett and Rawlings, 2004; Rawlings *et al.*, 2008). Members of this family differ from classical serine peptidases such as trypsin [Clan S (P) A] because they only hydrolyse peptides smaller than ~ 30 amino acids in length and hence has no known natural inhibitors (Morty *et al.*, 1999b; Morty *et al.*, 2001; Morty and Burleigh, 2004; Morty *et al.*, 2005a; Coetzer *et al.*, 2008). The peptidase has been studied as a target for diseases such as African trypanosomosis, Chagas disease and Leishmaniasis (Caler *et al.*, 1998; Coetzer *et al.*, 2008; Munday *et al.*, 2011). In kinetoplastid parasites, OPB is encoded for by a single copy gene and is expressed at all life cycle stages (Burleigh *et al.*, 1997; Morty *et al.*, 1999b; de Matos Guedes *et al.*, 2007). In the present study it was shown that OPB from *T. brucei* (*Tb*CATL) was down-regulated using RNAi. Results showed that parasites with up to an 87% reduction in expressed OPB were viable and able to divide at a rate equal to that of control parasites not induced for RNAi. Tubulin was targeted for degradation as a positive control for RNAi where round parasites that have lost their morphology (FAT cells) were observed (Inoue *et al.*, 2002; Coustou *et al.*, 2010). Parasites

induced for tubulin RNAi (+Tet) divided at a slightly lower rate when compared to parasites not targeted for tubulin RNA degradation. Tubulin is an essential structural protein in trypanosomes responsible for maintaining their physical morphology and involved in cytokenesis during cell division (Ngo *et al.*, 1998; Inoue *et al.*, 2002). The similar rate of division in parasites not induced for tubulin RNAi to those induced (Section 3.3.4, Fig 3.12; *Tco*TUB +/-Tet) can be attributed to the leakage of siRNA specific for tubulin that occurs even in the absence of tetracycline (Inoue *et al.*, 2002). It must also be taken into account that permanently transfected cells have only the two copies of the RNAi plasmid integrated into their genome, thus needing a longer period for the RNAi effect to be observed (Inoue *et al.*, 2002).

When compared to the tubulin control, RNAi targeting OPB in *T. congolense* was not successful (Section 3.3.4). RNAi was used as a tool in this study to assess the importance of OPB for parasite survival *in vitro*. The p2T7^{Ti} RNAi vector used for this study was originally constructed for RNAi studies in *T. brucei* (Wirtz *et al.*, 1999; Inoue *et al.*, 2002). An additional tetracycline repressor is required in RNAi inducible *T. congolense* strains (TRUM183:29-13) to control leakage in non-induced (-Tet) controls (Inoue *et al.*, 2002). The hairpin-loop RNAi vector (pZJM) developed for *T. brucei* (Wirtz *et al.*, 1999) has been successfully used for *in vivo* studies and has proven to be stable, inheritable and is tightly regulated using doxycycline (Abdulla *et al.*, 2008). The development of a stable hairpin-loop vector for *T. congolense* could alleviate the shortcomings associated with using the p2T7^{Ti} vector. RNAi targeting OPB in *T. brucei* was possible and knockdowns were assessed using western blotting, RT-PCR and RT-qPCR (Section 3.3.4). *T. brucei* parasites induced for OPB RNAi grew at a comparable rate to non-induced parasites and following these observations, it was concluded that *Tb*OPB is a non-essential peptidase *in vitro*. This indicated that OPB may have overlapping functions with other serine oligopeptidases within the parasite *in vitro*. It was therefore possible to generate *Tb*OPB null mutant parasites as described in Chapter 4.

The observations made on inducing *Tb*OPB RNAi *in vitro* correlate well with results observed when the *OPB* gene was knocked out in *L. major* (Munday *et al.*, 2011), *L. mexicana* (Pollock *et al.*, 2003) and *T. cruzi* (Caler *et al.*, 1998). *Leishmania* and *T. cruzi* are both intracellular parasites and knocking out the single copy gene encoding for OPB in all three cases does not affect parasite differentiation, morphology or motility *in vitro* and a phenotype is only observed during infection studies. *T. cruzi* ∆*opb* null mutants used in an *in vitro* assay to infect non-

phagocytic cells showed a 75% decline in the ability to infect normal rat kidney fibroblasts and L6E9 myoblasts (Caler *et al.*, 1998). Similar results were obtained when OPB was deleted in *Leishmania* where mutant parasites were significantly less able to infect and survive in macrophages (Munday *et al.*, 2011). *T. brucei* is an exclusively extracellular parasite but is able to cross endothelial barriers in the mammalian host and access the brain and testes (Grab *et al.*, 2004; Grab and Kennedy, 2008; Claes *et al.*, 2009; Grab *et al.*, 2009). The *T. brucei* ∆*opb* null mutants generated in the present study were able to cross the endothelial barrier in mouse testes indicating that this process occurs independent of *Tb*OPB and may instead be mediated by CATL peptidases in a calcium dependent reaction as has been shown in *T. b. gambiense* parasites (Nikolskaia *et al.*, 2006a; Nikolskaia *et al.*, 2006b). Previous experiments have shown that *Tb*OPB is inhibited by trypanocidal drugs (Morty *et al.*, 1998) and phosphonate diphenyl esters (Morty *et al.*, 2000). The identification of the peptidase irreversibly inhibited by phosphonate diphenyl esters using active site labelling with fluorescent inhibitors that have an Arg or Arg analogue residue in P_1, which are specific for *Tb*OPB, showed a band at 80 kDa, which is the predicted size of *Tb*OPB (Morty *et al.*, 2000). However, the fluorescent irreversible inhibitors also targeted other peptidases in *T. b. brucei* lysates, at 45 kDa, 68 kDa and 120 kDa. This would imply that more than one peptidase was inhibited leading to a loss of parasite viability and eventual death *in vitro* and during infection in mice. This was confirmed in the present study where *T. brucei Δopb* null mutants were successfully generated and able to infect mice with equal virulence to wild type parasites. Kinetoplastid serine peptidases play a variety of functions during infection in the mammalian host. Serine peptidase roles involved in host-parasite relationships include invasion of host cells, virulence, degradation of host peptides and matrix proteins, and the modulation of host immune responses (Coetzer *et al.*, 2008; de Matos Guedes *et al.*, 2010; Alvarez *et al.*, 2011). Some of these functions are summarised in Table 5.2 and discussed below.

Oligopeptidase B activity in *Leishmania* is highest during the amastigote stage in the host and is expressed at higher levels when compared to related parasites such as *T. cruzi* (Pollock *et al.*, 2003). In *L. donovani Δopb* null mutants, there is an accumulation of inactive enolase and plasminogen on the parasite cell surface which leads to an increase in macrophage gene transcription and up-regulated expression. Infected macrophages are thus activated leading to their clearance and ultimately resulting in reduced parasite virulence (Pollock *et al.*, 2003). *L. donovani* and *L. major Δopb* null mutants were also less able to infect host macrophages *in vitro*, but were as virulent in causing foot pad lesions *in vivo* in mice when compared to a wild

type infection (Pollock *et al.*, 2003; Munday *et al.*, 2011). It has been suggested that other oligopeptidase-B like peptidases might be responsible for compensating for the loss of OPB activity in *L. major Δopb* null mutants (Munday *et al.*, 2011). OPB2 in *L. amazonensis* is a related peptidase with a C-terminal extension that has similar reactive residues to classical OPB and could have overlapping roles (de Matos Guedes *et al.*, 2008). In *T. cruzi*, oligopeptidase B is crucial for the invasion of non-phagocytic cells where it activates a signalling pathway by mobilising Ca^{2+} stores in mammalian cells (Burleigh *et al.*, 1997; Caler *et al.*, 1998). *Tcr*OPB processes a precursor molecule that generates a Ca^{2+} agonist that is responsible for mobilising Ca^{2+} in the mammalian cell being invaded. The influx of Ca^{2+} regulates the attachment of host cell lysosomes to the membrane thus enabling parasite entry (Caler *et al.*, 1998).

The deletion of *Tb*OPB from *T. brucei* parasites resulted in a more than 5-fold increase in *Tb*POP-like activity (Chapter 4). In *T. brucei* and *T. cruzi*, prolyl oligopeptidase is encoded for by a single copy gene with 83% identity in amino acid sequences in the catalytic domain between the two parasites (Bastos *et al.*, 2005; Bastos *et al.*, 2010). In *T. cruzi*, *Tcr*POP has been implicated in assisting parasites gaining entry into murine muscle L-6 cells (Bastos *et al.*, 2005). This was observed after parasites, incubated with a specific irreversible chloromethane *Tcr*POP inhibitor showed a dose dependent reduction in the number of intracellular parasites (Bastos *et al.*, 2005). However, as with oligopeptidase B, this can only be confirmed using *Tcr*POP null mutant parasites. In *T. brucei*, POP is capable of hydrolysing native collagen and fibronectin which are rich in Gly-Pro residues (Bastos *et al.*, 2010) and thus could contribute to the presence of *T. brucei* parasites in the extravascular space and their ability to cross endothelial barriers in the brain and testes (Grab and Kennedy, 2008; Claes *et al.*, 2009; Bastos *et al.*, 2010).

It has recently been shown that some kinetoplastid serine proteases are also able to hydrolyse large substrates such as haemoglobin, BSA and ovalbumin (da Silva-López *et al.*, 2010). *L. chagasi* promastigotes were used to generate detergent (*Lc*CSI), aqueous (*Lc*CSII) and extracellular (*Lc*CSIII) fractions containing serine peptidase activity that could hydrolyse *N-ρ*-tosyl-L-arginine methyl ester (L-TAME) (da Silva-López *et al.*, 2010). *Lc*CSI and *Lc*CSII unlike *Lc*CSIII, are able to hydrolyse haemoglobin, BSA and ovalbumin and display maximal activity between pH 7 - 8.5.

Table 5.2 Summary of life stage expression, cellular distribution and functions of serine peptidases

Species	Peptidase Name(s)	Life stage(s) expressed	Cellular distribution	Assigned Functions	References
T. congolense	*Tco*OPB (OP-Tc) oligopeptidase B	All	Cytoplasm	Degradation of host peptides	i, ii
T. brucei	*Tb*OPB (OP-Tb) oligopeptidase B	All	Cytoplasm	Degradation of host peptides at Arg and Lys residues	ii, iii, iv, v
	*Tb*POP prolyl oligopeptidase	All	Cytoplasm	Degradation of host peptides at Pro residues, cleaves collagen I and IV	vi
L. major	*Lm*OPB oligopeptidase B	All, but highest in amastigotes	Cytoplasm	Invasion of host macrophages, virulence	vii
L.donovani	*Ld*OPB oligopeptidase B	All, but highest in amastigotes	Cytoplasm	Accumulation of enolase on parasite cell surface, immune modulation of infected macrophages, virulence	viii
L. chagasi	*Lc*CSI (detergent extract), *Lc*CSII (aqueous extract), *Lc*CSIII (extracellular)	Promastigote	Cytoplasm, flagellar pocket	Hydrolysis of haemoglobin, BSA and ovalbumin	ix
L. amazonensis	*La*SP (all serine peptidases)	All	Cytoplasm, flagellar pocket, whole cell lysate	Increase susceptibility to cutaneous leishmaniasis, activate IL-4 mediated Th2 response, and increase cutaneous hypersensitivity.	x, xi, xii, xiii
T. cruzi	*Tcr*OPB (Tc120) oligopeptidase B	All	Cytoplasm	Invasion of host non-phagocytic cells by mobilising mammalian intracellular Ca^{2+} stores	xiv, xv
	*Tcr*POP (Tc80) prolyl oligopeptidase	Amastigote, trypomastigote, epimastigote	Extracellular, para-flagellar vessicles	Degradation of host peptides at Pro residues, cleaves fibronectin, collagen I and IV, invasion of host non-phagocytic cells, virulence	xvi, xvii, xviii,
	*Tcr*SCP serine carboxypeptidase	Amastigote, epimastigote	Lysosome	Processed by *Tcr*CATL	xix

[i]Morty *et al.*(1999a), [ii]Coetzer *et al.* (2008), [iii]Troeberg *et al.* (1996b), [iv]Morty *et al.* (1999b), [v]Morty *et al.* (2001), [vi]Bastos *et al.* (2010), [vii]Munday *et al.* (2011), [viii]Swenerton *et al.* (2011), [ix]da Silva-López *et al.* (2010), [x]da Silva-López and De Simone (2004), [xi]de Matos Guedes *et al.* (2007), [xii]Santana *et al.* (1997), [xiii]de Matos Guedes *et al.* (2008), [xiv]de Matos Guedes *et al.* (2010), [xv]Burleigh *et al.* (1997), [xvi]Caler *et al.* (1998), [xvii]Grellier *et al.* (2001), [xviii]Bastos *et al.* (2005), [xix]Parussini *et al.* (2003)

A detergent soluble fraction with serine peptidase activity that is able to hydrolyse haemoglobin, BSA, ovalbumin and gelatin has also been isolated from *L. amazonensis* (da Silva-López and De Simone, 2004). Inhibition of *L. amazonensis* with N-tosyl-L-phenylalanine chloromethyl ketone (TPCK) and benzamidine serine peptidase inhibitors reduces parasite viability and causes lethal morphological alterations (Silva-Lopez *et al.*, 2007). Intradermal vaccination of mice with *L. amazonensis* promastigote antigens (LaAg) increases the susceptibility of infected mice to cutaneous leishmaniasis (Velez *et al.*, 2005). This susceptibility is reversed when mice are first vaccinated with a parasite extract that has been treated with serine peptidase inhibitors (de Matos Guedes *et al.*, 2010). It was observed that serine peptidases are responsible for activating an IL-4 mediated Th-2 response that does not release IFN-γ. This is similar to the scenario found in *L. mexicana* infections where *Lm*CATL null mutant parasites are unable to switch the host Th-1 response to a Th-2 response that is not as effective in controlling parasite burden (Alexander *et al.*, 1998).

A *T. cruzi* serine carboxypeptidase with an unknown function has been recombinantly expressed in the baculovirus expression system (Parussini *et al.*, 2003). Processing the propeptide of *T. cruzi* serine carboxypeptidase using *Tcr*CATL resulted in a mature protein with three-fold increase in activity, suggesting that some serine peptidases may also be substrates for kinetoplastid CATLs (Alvarez *et al.*, 2011).

This study has established through RNAi *in vivo* that *Tco*CATLs are important pathogenic factors that have an effect on parasite burden during infection in a mouse model. This finding is similar to observations on the role of *Tb*CATLs during infection in mice where disease progression is reduced when *Tb*CATLs are targeted for RNAi *in vivo* (Abdulla *et al.*, 2008). The development of a drug or vaccine that targets parasite CATLs would assist trypanosusceptible breeds of cattle to remain productive with lower parasite burden, reduced anaemia and inflammation associated with the disease. This study has also contributing to elucidating a role for *Tb*OPB and the overlapping functions it possibly shares with closely related *Tb*POP-like peptidases. When developing a drug targeting oligopeptidase B, it will be necessary to consider inhibitors that have an effect on more than one parasite oligopeptidase without affecting host peptidase functions. It is therefore important to characterise *Tb*OPB-like and *Tb*POP-like peptidases, and study their roles using RNAi or knock-out techniques. It must also be noted that although an overlap in function might occur between *Tb*OPB and related enzymes in the acute infection mouse model, a very different scenario may be seen

when using a chronic model for infection. Mice infected with *T.b. brucei Δopb* mutants do not survive the first peak of infection before *Tb*OPB is released by dying or dead parasites. Using larger mammalian hosts such as goats or cattle, where a natural chronic infection occurs, would provide new information on the role of OPB in trypanosomosis. The cyclical killing of parasites by the host immune system during the chronic phase of the disease would lead to an accumulation of OPB in control livestock infected with wild-type parasites and pathogenic effects such as the hydrolysis of host peptides could be compared to a livestock group infected with *T.b. brucei Δopb* parasites. Although this work suggests that *Tb*POP and *Tb*OPB may target the same host peptides, a chronic infection in livestock would elucidate the specific functions of each peptidase and identify which serine oligopeptidase is predominantly responsible for host peptide hydrolysis and consequent pathologies associated with the disease.

In conclusion, kinetoplastid parasite peptidases play various roles important for parasite viability within the host. Many of these peptidases seem to possess overlapping functions and are able to switch roles depending on the environment the parasite encounters. This has enabled trypanosomes to survive in different cells, tissues and hosts due to the function of different peptidases expressed at different stages of the life cycle. The loss in function of one peptidase can be compensated for by an up-regulation in activity of a closely related peptidase that could hydrolyse the same physiological substrate. This study has helped describe possible overlapping functions between peptidases involved in pathogenesis during trypanosome infection. Future areas of study should include using *T.b. brucei Δopb* parasites to infect ruminants where hormone dysregulation caused by parasite oligopeptidases can be adequately studied, and the generation of OPB/POP null mutants that would confirm the overlapping functions between these two peptidases. This information can be used when developing chemotherapeutic and vaccination strategies for trypanosomosis.

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 Neglected Tropical Diseases* **2**: e298.

Abrahamson, M., Alvarez-Fernandez, M. and Nathanson, C. M. (2003) Cystatins, *Biochemical Society Symposia* **70**: 179-199.

Aksoy, S., Maudlin, I., Dale, C., Robinson, A. S. and O'Neill, S. L. (2001) Prospects for control of African trypanosomiasis by tsetse vector manipulation, *Trends Parasitology* **17**: 29-35.

Aksoy, S. (2003) Control of tsetse flies and trypanosomes using molecular genetics, *Veterinary Parasitology* **115**: 125-145.

Alexander, J., Coombs, G. H. and Mottram, J. C. (1998) *Leishmania mexicana* cysteine proteinase-deficient mutants have attenuated virulence for mice and potentiate a Th1 response, *Journal of Immunology* **161**: 6794-6801.

Alsford, S., Turner, D. J., Obado, S. O., Sanchez-Flores, A., Glover, L., Berriman, M., Hertz-Fowler, C. and Horn, D. (2011) High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome, *Genome Research* **21**: 915-924.

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

Andrews, N. W. and Whitlow, M. B. (1989) Secretion by *Trypanosoma cruzi* of a hemolysin active at low pH, *Molecular and Biochemical Parasitology* **33**: 249-256.

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

Antoine-Moussiaux, N., Büscher, P. and Desmecht, D. (2009) Host-Parasite Interactions in Trypanosomiasis: on the Way to an Antidisease Strategy, *Infection and Immunity* **77**: 1276-1284.

Ariyanayagam, M. R., Oza, S. L., Guther, M. L. and Fairlamb, A. H. (2005) Phenotypic analysis of trypanothione synthetase knockdown in the African trypanosome, *Biochemical Journal* **391**: 425-432.

Ashmun, R. A. and Look, A. T. (1990) Metalloprotease activity of CD13/aminopeptidase N on the surface of human myeloid cells, *Blood* **75**: 462-469.

Askonas, B. A. (1985) Macrophages as mediators of immunosuppression in murine African trypanosomiasis, *Current Topics in Microbiology and Immunology* **117**: 119-127.

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, *Molecular and Biochemical Parasitology* **56**: 103-116.

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 Immunology* **15**: 465-474.

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 Immunology* **15**: 101-111.

Authié, E. (1994) Trypanosomiasis and trypanotolerance in cattle: a role for congopain?, *Parasitology 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, *International Journal of Parasitolgy* **31**: 1429-1433.

Baetselier, P. D., Namangala, B., Noel, W., Brys, L., Pays, E. and Beschin, A. (2001) Alternative versus classical macrophage activation during experimental African trypanosomosis, *International journal for parasitology* **31**: 575-587.

Balaña-Fouce, R. and Reguera, R. M. (2007) RNA interference in *Trypanosoma brucei*: a high-throughput engine for functional genomics in trypanosomatids?, *Trends in parasitology* **23**: 348-351.

Bangs, J. D., Ransom, D. A., Nimick, M., Christie, G. and Hooper, N. M. (2001) *In vitro* cytocidal effects on *Trypanosoma brucei* and inhibition of *Leishmania major* GP63 by peptidomimetic metalloprotease inhibitors, *Molecular and Biochemical Parasitology* **114**: 111-117.

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

Barnes, R. L. and McCulloch, R. (2007) *Trypanosoma brucei* homologous recombination is dependent on substrate length and homology, though displays a differential dependence on mismatch repair as substrate length decreases, *Nucleic Acids Research* **35**: 3478-3493.

Barrett, A. J. and McDonald, J. K. (1986) Nomenclature: protease, proteinase and peptidase, *Biochemical Journal* **237**: 935.

Barrett, A. J. and Rawlings, N. D. (2004) The clans and families of cysteine peptidases. In: *Handbook of proteolytic enzymes* Barrett, A. J., Rawlings, N. D. and Woessner, J. F. (eds.), Elsevier, London. 1160-1163.

Barrett, M. P., Burchmore, R. J., Stich, A., Lazzari, J. O., Frasch, A. C., Cazzulo, J. J. and Krishna, S. (2003) The trypanosomiases, *Lancet* **362**: 1469-1480.

Bastos, I. M. D., Grellier, P., Martins, N. F., Cadavid-Restrepo, G., de Souza-Ault, M. R., Augustyns, K., Teixeira, A. R. L., Schrével, J., Maigret, B., da Silveira, J. F. and Santana, J. M. (2005) Molecular, functional and structural properties of the prolyl oligopeptidase of *Trypanosoma cruzi* (POP Tc80), which is required for parasite entry into mammalian cells, *Biochemical Journal* **388**: 29-38.

Bastos, I. M. D., Motta, F. N., Charneau, S., Santana, J. M., Dubost, L., Augustyns, K. and Grellier, P. (2010) Prolyl oligopeptidase of *Trypanosoma brucei* hydrolyzes native collagen, peptide hormones and is active in the plasma of infected mice, *Microbes and Infection* **12**: 457-466.

Baum, J., Papenfuss, A. T., Mair, G. R., Janse, C. J., Vlachou, D., Waters, A. P., Cowman, A. F., Crabb, B. S. and de Koning-Ward, T. F. (2009) Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites, *Nucleic Acids Research* **37**: 3788-3798.

Baumeister, R. (2002) Cross-species studies for target validation, *Briefings in Funtional Genomics* **1**: 53-65.

Bellofatto, V. and Palenchar, J. B. (2008) RNA Interference as a Genetic Tool in Trypanosomes. In: Barik, S. (eds.), Humana Press, New York. 83-94.

Berg, M., Van der Veken, P., Joossens, J., Muthusamy, V., Breugelmans, M., Moss, C. X., Rudolf, J., Cos, P., Coombs, G. H., Maes, L., Haemers, A., Mottram, J. C. and Augustyns, K. (2010) Design and evaluation of *Trypanosoma brucei* metacaspase inhibitors, *Bioorganic and Medicinal Chemistry Letters* **20**: 2001-2006.

Berger, A. and Schechter, I. (1970) Mapping the active site of papain with the aid of peptide substrates and inhibitors, *Philosophical Transactions of the Royal Society B: Biological Sciences* **257**: 249-264.

Berriman, M. Ghedin, E. Hertz-Fowler, C. Blandin, G. Renauld, H. Bartholomeu, D. C. Lennard, N. J. Caler, E. Hamlin, N. E. Haas, B. Bohme, U. Hannick, L. Aslett, M. A. Shallom, J. Marcello, L. Hou, L. Wickstead, B. Alsmark, U. C. Arrowsmith, C. Atkin, R. J. Barron, A. J. Bringaud, F. Brooks, K. Carrington, M. Cherevach, I. Chillingworth, T. J. Churcher, C. Clark, L. N. Corton, C. H. Cronin, A. Davies, R. M. Doggett, J. Djikeng, A. Feldblyum, T. Field, M. C. Fraser, A. Goodhead, I. Hance, Z. Harper, D. Harris, B. R. Hauser, H. Hostetler, J. Ivens, A. Jagels, K. Johnson, D. Johnson, J. Jones, K. Kerhornou, A. X. Koo, H. Larke, N. Landfear, S. Larkin, C. Leech, V. Line, A. Lord, A. Macleod, A. Mooney, P. J. Moule, S. Martin, D. M. Morgan, G. W. Mungall, K. Norbertczak, H. Ormond, D. Pai, G. Peacock, C. S. Peterson, J. Quail, M. A. Rabbinowitsch, E. Rajandream, M. A. Reitter, C. Salzberg, S. L. Sanders, M. Schobel, S. Sharp, S. Simmonds, M. Simpson, A. J. Tallon, L. Turner, C. M. Tait, A. Tivey, A. R. Van Aken, S. Walker, D. Wanless, D. Wang, S. White, B. White, O. Whitehead, S. Woodward, J. Wortman, J. Adams, M. D. Embley, T. M. Gull, K. Ullu, E. Barry, J. D. Fairlamb, A. H. Opperdoes, F. Barrell, B. G. Donelson, J. E. Hall, N. Fraser, C. M. Melville, S. E. and El-Sayed, N. M. (2005) The genome of the African trypanosome *Trypanosoma brucei*, *Science* **309**: 416-422.

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.

Black, S. J., Sicard, E. L., Murphy, N. and Nolan, D. (2001) Innate and acquired control of trypanosome parasitaemia in Cape buffalo, *International Journal of Parasitolgy* **31**: 562-565.

Blackman, M. J. (2003) RNAi in protozoan parasites: what hope for the Apicomplexa?, *Protist* **154**: 177-180.

Bonetta, L. (2004) RNAi: Silencing never sounded better, *Nat Meth* **1**: 79-86.

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*, *International journal for parasitology* **31**: 1435-1440.

Boulangé, A., 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 Expression and Purification* **75**: 95-103.

Brady, C. P., Dowd, A. J., Brindley, P. J., Ryan, T., Day, S. R. and Dalton, J. P. (1999) Recombinant Expression and Localization of *Schistosoma mansoni* Cathepsin L1 Support Its Role in the Degradation of Host Hemoglobin, *Infection and Immunity* **67**: 368-374.

Brew, K., Dinakarpandian, D. and Nagase, H. (2000) Tissue inhibitors of metalloproteinases: evolution, structure and function, *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1477**: 267-283.

Brown, K. (2008) From Ubombo to Mkhuzi: Disease, Colonial Science, and the Control of *Nagana* (Livestock Trypanosomosis) in Zululand, South Africa, c. 1894-1953, *Journal of The History of Medicine and Allied Sciences* **63**: 285-322.

Burkard, G., Fragoso, C. M. and Roditi, I. (2007) Highly efficient stable transformation of bloodstream forms of *Trypanosoma brucei*, *Molecular and Biochemical Parasitology* **153**: 220-223.

Burleigh, B. A., Caler, E. V., Webster, P. and Andrews, N. W. (1997) A cytosolic serine endopeptidase from *Trypanosoma cruzi* is required for the generation of Ca^{2+} signaling in mammalian cells, *Journal of Cell Biology* **136**: 609-620.

Büscher, G. and Friedhoff, K. T. (1984) The morphology of ovine *Trypanosoma melophagium* (zoomastigophorea: kinetoplastida), *Journal of Protozoology* **31**: 98-101.

Cadavid-Restrepo, G., Gastardelo, T., Faudry, E., de Almeida, H., Bastos, I., Negreiros, R., Lima, M., Assumpcao, T., Almeida, K., Ragno, M., Ebel, C., Ribeiro, B., Felix, C. and Santana, J. (2011) The major leucyl aminopeptidase of *Trypanosoma cruzi* (LAPTc) assembles into a homohexamer and belongs to the M17 family of metallopeptidases, *BMC Biochemistry* **12**: 46.

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*, *Molecular and Biochemical Parasitology* **118**: 61-73.

Caffrey, C. R. and Steverding, D. (2009) Kinetoplastid papain-like cysteine peptidases, *Molecular and Biochemical Parasitology* **167**: 12-19.

Caffrey, C. R., Lima, A.-P. and Steverding, D. (2011) Cysteine Peptidases of Kinetoplastid Parasites, *Advances in Experimental Medicine and Biology* **712**: 84-99.

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

Campetella, O., Henriksson, J., Aslund, L., Frasch, A. C., Pettersson, U. and Cazzulo, J. J. (1992) The major cysteine proteinase (cruzipain) from *Trypanosoma cruzi* is encoded by multiple polymorphic tandemly organized genes located on different chromosomes, *Molecular and Biochemical Parasitology* **50**: 225-234.

Carruthers, V. B., van der Ploeg, L. H. and Cross, G. A. (1993) DNA-mediated transformation of bloodstream-form *Trypanosoma brucei*, *Nucleic Acids Research* **21**: 2537-2538.

Cazzulo, J. J. and Frasch, A. C. (1992) SAPA/trans-sialidase and cruzipain: two antigens from *Trypanosoma cruzi* contain immunodominant but enzymatically inactive domains, *FASEB Journal* **6**: 3259-3264.

Cerutti, L., Mian, N. and Bateman, A. (2000) Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain, *Trends in Biochemical Sciences* **25**: 481-482.

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*, *Molecular and Biochemical Parasitology* **88**: 85-94.

Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction, *Analytical Biochemistry* **162**: 156-159.

Claes, F., Vodnala, S. K., Van Reet, N., Boucher, N., Lunden-Miguel, H., Baltz, T., Goddeeris, B. M., Büscher, P. and Rottenberg, M. E. (2009) Biolumnescent imaging of *Trypanosoma brucei* shows preferential testis dissemination which may hamper drug efficacy in sleeping sickness, *PLoS Neglected Tropical Diseases* **3**: e486.

Clayton, C. E. (1999) Genetic manipulation of kinetoplastida, *Parasitology Today* **15**: 372-378.

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

Coustou, V., Guegan, F., Plazolles, N. and Baltz, T. (2010) Complete in vitro life cycle of *Trypanosoma congolense*: development of genetic tools, *PLoS Neglected Tropical Diseases* **4**: e618.

Coustou, V., Plazolles, N., Guegan, F. and Baltz, T. (2011) Sialidases play a key role in infection and anaemia in *Trypanosoma congolense* animal trypanosomiasis, *Cellular Microbiology*: In Press.

Cox, M. M. (2003) The bacterial RecA protein as a motor protein, *Annual Review of Microbiology* **57**: 551-577.

Cruz, A., Coburn, C. M. and Beverley, S. M. (1991) Double targeted gene replacement for creating null mutants, *Proceedings of the National Academy of Sciences of the United States of America* **88**: 7170-7174.

Cruz, A. K., Titus, R. and Beverley, S. M. (1993) Plasticity in chromosome number and testing of essential genes in *Leishmania* by targeting, *Proceedings of the National Academy of Sciences of the United States of America* **90**: 1599-1603.

da Silva-López, R., dos Santos, T., Morgado-Díaz, J., Tanaka, M. and de Simone, S. (2010) Serine protease activities in *Leishmania (Leishmania) chagasi* promastigotes, *Parasitology Research* **107**: 1151-1162.

da Silva-López, R. E. and De Simone, S. G. (2004) A serine protease from a detergent-soluble extract of *Leishmania (Leishmania) amazonensis*, *Z Naturforsch C* **59**: 590-598.

Damasceno, J., Beverley, S. and Tosi, L. (2010) A transposon toolkit for gene transfer and mutagenesis in protozoan parasites, *Genetica* **138**: 301-311.

DaRocha, W. D., Otsu, K., Teixeira, S. M. and Donelson, J. E. (2004) Tests of cytoplasmic RNA interference (RNAi) and construction of a tetracycline-inducible T7 promoter system in *Trypanosoma cruzi*, *Molecular and Biochemical Parasitology* **133**: 175-186.

de Matos Guedes, H., Carneiro, M., Gomes, D., Rossi-Bergmanmn, B. and de Simone, S. (2007) Oligopeptidase B from *L. amazonensis*: molecular cloning, gene expression analysis and molecular model, *Parasitology Research* **101**: 853-863.

de Matos Guedes, H., de Carvalho, R., de Oliveira Gomes, D., Rossi-Bergmann, B. and De-Simone, S. (2008) Oligopeptidase B-2 from *Leishmania amazonensis* with an unusual C-terminal extension, *Acta Parasitologica* **53**: 197-204.

de Matos Guedes, H. L., Pinheiro, R. O., Chaves, S. P., De-Simone, S. G. and Rossi-Bergmann, B. (2010) Serine proteases of *Leishmania amazonensis* as immunomodulatory and disease-aggravating components of the crude LaAg vaccine, *Vaccine* **28**: 5491-5496.

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

Deutscher, D., Meilijson, I., Schuster, S. and Ruppin, E. (2008) Can single knockouts accurately single out gene functions?, *BioMed Central Systems Biology* **2**: 50.

Durand-Dubief, M. and Bastin, P. (2003) TbAGO1, an argonaute protein required for RNA interference, is involved in mitosis and chromosome segregation in *Trypanosoma brucei*, *BioMed Central Biology* **1**: 2.

El-Sayed, N. M., Hedge, P., Quakenbush, J., Melville, S. E. and Donelson, J. E. (2000) The African trypanosome genome, *International Journal of Parasitolgy* **30**: 329-345.

Estevez, A. M., Kierszenbaum, F., Wirtz, E., Bringaud, F., Grunstein, J. and Simpson, L. (1999) Knockout of the glutamate dehydrogenase gene in bloodstream *Trypanosoma brucei* in culture has no effect on editing of mitochondrial mRNAs, *Molecular and Biochemical Parasitology* **100**: 5-17.

Ferreira, V., Molina, M. C., Schwaeble, W., Lemus, D. and Ferreira, A. (2005) Does *Trypanosoma cruzi* calreticulin modulate the complement system and angiogenesis?, *Trends in parasitology* **21**: 169-174.

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

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.

Frame, M. J., Mottram, J. C. and Coombs, G. H. (2000) Analysis of the roles of cysteine proteinases of *Leishmania mexicana* in the host-parasite interaction, *Parasitology* **121 (Pt 4)**: 367-377.

Freedman, D. J. and Beverley, S. M. (1993) Two more independent selectable markers for stable transfection of *Leishmania*, *Molecular and Biochemical Parasitology* **62**: 37-44.

Fülöp, V., Bocskei, Z. and Polgar, L. (1998) Prolyl oligopeptidase: an unusual beta-propeller domain regulates proteolysis, *Cell* **94**: 161-170.

Gamboa de Dominguez, N. and Rosenthal, P. (1996) Cysteine proteinase inhibitors block early steps in hemoglobin degradation by cultured malaria parasites, *Blood* **87**: 4448-4454.

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, *Journal of Biological Chemistry* **281**: 1741-1745.

Gaud, A., Carrington, M., Deshusses, J. and Schaller, D. R. (1997) Polymerase chain reaction-based gene disruption in *Trypanosoma brucei*, *Molecular and Biochemical Parasitology* **87**: 113-115.

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

Gerczei, T., Keseru, G. M. and Naray-Szabo, G. (2000) Construction of a 3D model of oligopeptidase B, a potential processing enzyme in prokaryotes, *Journal of Molecular Graphics and Modelling* **18**: 7-17, 57-18.

Gibson, W. C., Lom, J., Peckova, H., Ferris, V. R. and Hamilton, P. B. (2005) Phylogenetic analysis of freshwater fish trypanosomes from Europe using ssu rRNA gene sequences and random amplification of polymorphic DNA, *Parasitology* **130**: 405-412.

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

Grab, D. J., Nikolskaia, O. V., Kim, Y. V., Lonsdale-Eccles, J. D., Ito, S., Hara, T., Fukuma, T., Nyarko, E., kim, K. J. and Stins, M. F. (2004) African trypanosome interactions with an *in vitro* model of the human blood-brain barrier, *Journal of Parasitology* **90**: 970-979.

Grab, D. J. and Kennedy, P. G. (2008) Traversal of human and animal trypanosomes across the blood-brain barrier, *Journal for NeuroVirology* **14**: 344-351.

Grab, D. J., Garcia-Garcia, J. C., Nikolskaia, O. V., Kim, Y. V., Brown, A., Pardo, C. A., Zhang, Y., Becker, K. G., Wilson, B. A., de, A. L. A. P., Scharfstein, J. and Dumler, J. S. (2009) Protease activated receptor signaling is required for african trypanosome traversal of human brain microvascular endothelial cells, *PLoS Neglected Tropical Diseases* **3**: e479.

Grellier, P., Vendeville, S., Joyeau, R., Bastos, I. M., Drobecq, H., Frappier, F., Teixeira, A. R., Schrevel, J., Davioud-Charvet, E., Sergheraert, C. and Santana, J. M. (2001) *Trypanosoma cruzi* prolyl oligopeptidase Tc80 is involved in nonphagocytic mammalian cell invasion by trypomastigotes, *Journal of Biological Chemistry* **276**: 47078-47086.

Grisard, E. C. (2002) Salivaria or Stercoraria? The *Trypanosoma rangeli* dilemma, *Kinetoplastid Biology and Disease* **1**: 5.

Guo, S. and Kemphues, K. J. (1995) par-1, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed., *Cell* **81**: 611-620.

Gutierrez, R. J. (1989) Hematozoa from the spotted owl, *Journal of Wildlife Diseases* **25**: 614-618.

Hajduk, S. and Vickerman, K. (1981) Antigenic differentiation of *Trypanosoma brucei*: studies on metacyclic and first parasitaemia populations, *Transactions of the Royal Society of Tropical Medicine and Hygiene* **75**: 145-146.

Hao, Z., Kasumba, I., Lehane, M. J., Gibson, W. C., Kwon, J. and Aksoy, S. (2001) Tsetse immune responses and trypanosome transmission: implications for the development of tsetse-based strategies to reduce trypanosomiasis, *Proceedings of the National Academy of Sciences of the United States of America* **98**: 12648-12653.

Harris, P. K., Yeoh, S., Dluzewski, A. R., O'Donnell, R. A., Withers-Martinez, C., Hackett, F., Bannister, L. H., Mitchell, G. H. and Blackman, M. J. (2005) Molecular identification of a malaria merozoite surface sheddase, *PLoS Pathogens* **1**: 241-251.

Hartley, C. L. and McCulloch, R. (2008) *Trypanosoma brucei* BRCA2 acts in antigenic variation and has undergone a recent expansion in BRC repeat number that is important during homologous recombination, *Molecular Microbiology* **68**: 1237-1251.

Heinzel, F. P., Schoenhaut, D. S., Rerko, R. M., Rosser, L. E. and Gately, M. K. (1993) Recombinant interleukin 12 cures mice infected with *Leishmania major*, *Journal of Experimental Medicine* **177**: 1505-1509.

Hilley, J. D., Zawadzki, J. L., McConville, M. J., Coombs, G. H. and Mottram, J. C. (2000) *Leishmania mexicana* mutants lacking glycosylphosphatidylinositol (GPI):protein transamidase provide insights into the biosynthesis and functions of GPI-anchored proteins, *Molecular Biology of the Cell* **11**: 1183-1195.

Hirumi, H. and Hirumi, K. (1989) Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers, *Journal of Parasitology* **75**: 985-989.

Hoare, C. A. (1972) The Trypanosomes of Mammals, *Journal of Small Animal Practice* **13**: 671-672.

Huson, L. E. (2006) Antibody-mediated inhibition of proteases of African trypanosomes. School of Biochemistry, Genetics and Microbiology. University of KwaZulu-Natal, Pietermaritzburg.

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 α_2-macroglobulin, *Veterinary Research* **40**: 52.

Hysek, J. and Zizka, Z. (1976) Transmission of *Trypanosoma rotatorium* from frogs to white mice, *Nature* **260**: 608-609.

Inoue, N., Otsu, K., Ferraro, D. M. and Donelson, J. E. (2002) Tetracycline-regulated RNA interference in *Trypanosoma congolense*, *Molecular and Biochemical Parasitology* **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, *Journal of Biological Chemistry* **281**: 33664-33676.

Jameel M, I. (2004) Parasite interaction with host complement: beyond attack regulation, *Trends in parasitology* **20**: 407-412.

Jaye, A. B., Nantulya, V. M., Majiwa, P. A., Urakawa, T., Masake, R. A., Wells, C. W. and ole-MoiYoi, O. K. (1994) A *Trypanosoma (Nannomonas) congolense*-specific antigen released into the circulation of infected animals is a thiol protease precursor *EMBL accession no. L25130*.

Jefferies, D., Tebabi, P., Ray, D. L. and Pays, E. (1993) The ble resistance gene as a new selectable marker for *Trypanosoma brucei*: fly transmission of stable procyclic transformants to produce antibiotic resistant bloodstream forms, *Nucleic Acids Research* **21**: 191-195.

Jirku, M., Kolesnikov, A. A., Benada, O. and Lukes, J. (1995) Marine fish and ray trypanosomes have large kinetoplast minicircle DNA, *Molecular and Biochemical Parasitology* **73**: 279-283.

Joshi, P. B., Webb, J. R., Davies, J. E. and McMaster, W. R. (1995) The gene encoding streptothricin acetyltransferase (sat) as a selectable marker for *Leishmania* expression vectors, *Gene* **156**: 145-149.

Joshi, P. B., Kelly, B. L., Kamhawi, S., Sacks, D. L. and McMaster, W. R. (2002) Targeted gene deletion in *Leishmania major* identifies leishmanolysin (GP63) as a virulence factor, *Molecular and Biochemical Parasitology* **120**: 33-40.

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

Kanatani, A., Masuda, T., Shimoda, T., Misoka, F., Lin, X. S., Yoshimoto, T. and Tsuru, D. (1991) Protease II from *Escherichia coli*: sequencing and expression of the enzyme gene and characterization of the expressed enzyme, *Journal of Biochemistry* **110**: 315-320.

Kangethe, R. T., Boulangé, A. F. V., Coustou, V., Baltz, T. and Coetzer, T. H. T. (2011) *Trypanosoma brucei brucei* oligopeptidase B null mutants display increased prolyl oligopeptidase-like activity, *Molecular and Biochemical Parasitology*, doi:10.1016/j.molbiopara.2011.11.007.

Kanti Bhattacharyya, M., Norris, D. E. and Kumar, N. (2004) Molecular players of homologous recombination in protozoan parasites: implications for generating antigenic variation, *Infection, Genetics and Evolution* **4**: 91-98.

Kappmeier, K., Nevill, E. M. and Bagnall, R. J. (1998) Review of tsetse flies and trypanosomosis in South Africa, *Ondersterpoort journal of veterinary researsch* **65**: 195-203.

Kinyua, J. K., Nguu, E. K., Mulaa, F. and Ndung'u, J. M. (2005) Immunization of rabbits with *Glossina pallidipes* tsetse fly midgut proteins: effects on the fly and trypanosome transmission, *Vaccine* **23**: 3824-3828.

Klemba, M., Gluzman, I. and Goldberg, D. E. (2004) A *Plasmodium falciparum* dipeptidyl aminopeptidase I participates in vacuolar hemoglobin degradation, *Journal of Biological Chemistry* **279**: 43000-43007.

Knowles, G., Black, S. J. and Whitelaw, D. D. (1987) Peptidase in the plasma of mice infected with *Trypanosoma brucei brucei*, *Parasitology* **95 (Pt 2)**: 291-300.

Kobayashi, K., Kamakura, T., Tanaka, T., Yamaguchi, I. and Endo, T. (1991) Nucleotide sequence of the bsr gene and N-terminal amino acid sequence of blasticidin S deaminase from blasticidin S resistant *Escherichia coli* TK121, *Agricultural and Biological Chemistry* **55**: 3155-3157.

Kohl, L. and Bastin, P. (2005) The flagellum of trypanosomes, *International Review of Cytology* **244**: 5336-5346.

Kristjanson, P. M., Swallow, B. M., Rowlands, G. J., Kruska, R. L. and de Leeuw, P. N. (1999) Measuring the costs of African animal trypanosomosis, the potential benefits of control and returns to research, *Agricultural Systems* **59**: 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, *Infection and Immunity* **75**: 2026-2034.

Kuzoe, F. A. S. (1993) Current situation of African trypanosomiasis, *Acta Tropica* **54**: 153-162.

Laban, A., Tobin, J. F., de Lafaille, M. A. C. and Wirth, D. F. (1990) Stable expression of the bacterial neo^{r} gene in *Leishmania enriettii*, *Nature* **343**: 572-574.

LaCount, D. J., Barrett, B. and Donelson, J. E. (2002) *Trypanosoma brucei* FLA1 is required for flagellum attachment and cytokinesis, *Journal of Biological Chemistry* **277**: 17580-17588.

LaCount, D. J., Gruszynski, A. E., Grandgenett, P. M., Bangs, J. D. and Donelson, J. E. (2003) Expression and function of the *Trypanosoma brucei* major surface protease (GP63) genes, *Journal of Biological Chemistry* **278**: 24658-24664.

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

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, *Journal of Biological Chemistry* **273**: 25112-25116.

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, *Biological Chemistry* **383**: 739-749.

Leal, S., Acosta-Serrano, A., Morris, J. and Cross, G. A. M. (2004) Transposon Mutagenesis of *Trypanosoma brucei* Identifies Glycosylation Mutants Resistant to Concanavalin A, *Journal of Biological Chemistry* **279**: 28979-28988.

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

Lecordier, L., Walgraffe, D., Devaux, S., Poelvoorde, P., Pays, E. and Vanhamme, L. (2005) *Trypanosoma brucei* RNA interference in the mammalian host, *Molecular and Biochemical Parasitology* **140**: 127-131.

Lee, M. G. and Van der Ploeg, L. H. (1990) Homologous recombination and stable transfection in the parasitic protozoan *Trypanosoma brucei*, *Science* **250**: 1583-1587.

Lee, M. G. and van der Ploeg, L. H. (1991) The hygromycin B-resistance-encoding gene as a selectable marker for stable transformation of *Trypanosoma brucei*, *Gene* **105**: 255-257.

Li, S.-Q., Yang, W.-B., Lun, Z.-R., Ma, L.-J., Xi, S.-M., Chen, Q.-L., Song, X.-W., Kang, J. and Yang, L.-Z. (2009) Immunization with recombinant actin from *Trypanosoma evansi* induces protective immunity against *T. evansi*, *T. equiperdum* and *T. b. brucei* infection, *Parasitology Research* **104**: 429-435.

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 Immunology* **29**: 191-199.

Lillico, S., Field, M. C., Blundell, P., Coombs, G. H. and Mottram, J. C. (2003) Essential roles for GPI-anchored proteins in African trypanosomes revealed using mutants deficient in GPI8, *Molecular Biology of the Cell* **14**: 1182-1194.

Lima, A. P., dos Reis, F. C., Serveau, C., Lalmanach, G., Juliano, L., Menard, 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, *Molecular and Biochemical Parasitology* **114**: 41-52.

Lima, A. P., Almeida, P. C., Tersariol, I. L., Schmitz, V., Schmaier, A. H., Juliano, L., Hirata, I. Y., Muller-Esterl, W., Chagas, J. R. and Scharfstein, J. (2002) Heparan sulfate modulates kinin release by *Trypanosoma cruzi* through the activity of cruzipain, *Journal of Biological Chemistry* **277**: 5875-5881.

Livak, K. J. and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, *Methods* **25**: 402-408.

Lonsdale-Eccles, J. D. and Grab, D. J. (2002) Trypanosome hydrolases and the blood-brain barrier, *Trends Parasitology* **18**: 17-19.

Lorenz, P., Maier, A. G., Baumgart, E., Erdmann, R. and Clayton, C. (1998) Elongation and clustering of glycosomes in *Trypanosoma brucei* overexpressing the glycosomal Pex11p, *EMBO Journal* **17**: 3542-3555.

Losos, G. J. and Ikede, B. O. (1972) Review of pathology of diseases in domestic and laboratory animals caused by *Trypanosoma congolense, T. vivax, T. brucei, T. rhodesiense* and *T. gambiense*, *Veterinary Pathology* **9**: 1-79.

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.

Luhrs, K. A., Fouts, D. L. and Manning, J. E. (2003) Immunization with recombinant paraflagellar rod protein induces protective immunity against *Trypanosoma cruzi* infection, *Vaccine* **21**: 3058-3069.

Lund, L. R., Romer, J., Bugge, T. H., Nielsen, B. S., Frandsen, T. L., Degen, J. L., Stephens, R. W. and Dano, K. (1999) Functional overlap between two classes of matrix-degrading proteases in wound healing, *EMBO Journal* **18**: 4645-4656.

Lye, L. F., Owens, K., Shi, H., Murta, S. M., Vieira, A. C., Turco, S. J., Tschudi, C., Ullu, E. and Beverley, S. M. (2010) Retention and loss of RNA interference pathways in trypanosomatid protozoans, *PLoS Pathogens* **6**: e1001161.

Mackey, Z. B., O'Brien, T. C., Greenbaum, D. C., Blank, R. B. and McKerrow, J. H. (2004) A cathepsin B-like protease is required for host protein degradation in *Trypanosoma brucei*, *Journal of Biological Chemistry* **279**: 48426-48433.

Magez, S., Stijlemans, B., Baral, T. and De Baetselier, P. (2002) VSG-GPI anchors of African trypanosomes: their role in macrophage activation and induction of infection-associated immunopathology, *Microbes and Infection* **4**: 999-1006.

Magez, S., Schwegmann, A., Atkinson, R., Claes, F., Drennan, M., De Baetselier, P. and Brombacher, F. (2008) The role of B-cells and IgM antibodies in parasitemia, anemia, and VSG switching in *Trypanosoma brucei*-infected mice, *PLoS Pathogens* **4**: e1000122.

Magez, S. and Radwanska, M. (2009) African trypanosomiasis and antibodies: implications for vaccination, therapy and diagnosis, *Future Microbiology* **4**: 1075-1087.

Magez, S., Caljon, G., Tran, T., Stijlemans, B. and Radwanska, M. (2010) Current status of vaccination against African trypanosomiasis, *Parasitology* **137**: 2017-2027.

Magez, S. and Caljon, G. (2011) Mouse models for pathogenic African trypanosomes: unravelling the immunology of host–parasite–vector interactions, *Parasite Immunology* **33**: 423-429.

Malhotra, P., Dasaradhi, P. V. N., Kumar, A., Mohmmed, A., Agrawal, N., Bhatnagar, R. K. and Chauhan, V. S. (2002) Double-stranded RNA-mediated gene silencing of cysteine proteases (falcipain-1 and -2) of *Plasmodium falciparum*, *Molecular Microbiology* **45**: 1245-1254.

Mamabolo, M. V., Ntantiso, L., Latif, A. and Majiwa, P. A. (2009) Natural infection of cattle and tsetse flies in South Africa with two genotypic groups of *Trypanosoma congolense*, *Parasitology* **136**: 425-431.

Martinez, J., Campetella, O., Frasch, A. C. and Cazzulo, J. J. (1991) The major cysteine proteinase (cruzipain) from *Trypanosoma cruzi* is antigenic in human infections, *Infection and Immunity* **59**: 4275-4277.

Masocha, W., Robertson, B., Rottenberg, M. E., Mhlanga, J. D., Sorokin, L. and Kristensson, K. (2004) Cerebral vessel laminins and IFN-γ define *Trypanosoma brucei brucei* penetration of the blood-brain barrier, *Journal of Clinical Investigation* **114**: 689-694.

Masocha, W., Rottenberg, M. E. and kristensson, K. (2006) Minocycline impedes African trypanosome invasion of the brain in a murine model, *Antimicrobial Agents and Chemotherapy* **50**: 1798-1804.

Masson, J. Y. and West, S. C. (2001) The Rad51 and Dmc1 recombinases: a non-identical twin relationship, *Trends in Biochemical Sciences* **26**: 131-136.

Matovu, E., Stewart, M. L., Geiser, F., Brun, R., Maser, P., Wallace, L. J., Burchmore, R. J., Enyaru, J. C., Barrett, M. P., Kaminsky, R., Seebeck, T. and de Koning, H. P. (2003) Mechanisms of arsenical and diamidine uptake and resistance in *Trypanosoma brucei*, *Eukaryotic Cell* **2**: 1003-1008.

Mbawa, Z. R., Gumm, I. D., Fish, W. R. and Lonsdale-Eccles, J. D. (1991) Endopeptidase variations among different life-cycle stages of African trypanosomes, *European Journal of Biochemistry* **195**: 183-190.

Mbawa, Z. R., Webster, P. and Lonsdale-Eccles, J. D. (1991b) Immunolocalization of a cysteine protease within the lysosomal system of *Trypanosoma congolense*, *European Journal of Cell Biology* **56**: 243-250.

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*, *European Journal of Biochemistry* **204**: 371-379.

McCulloch, R. and Barry, J. D. (1999) A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation, *Genes and Development* **13**: 2875-2888.

McCulloch, R. and Horn, D. (2009) What has DNA sequencing revealed about the VSG expression sites of African trypanosomes?, *Trends in parasitology* **25**: 359-363.

McDermott, J. J. and Coleman, P. G. (2001) Comparing apples and oranges-model-based assessment of different tsetse-transmitted trypanosomosis control strategies, *International journal for parasitology* **31**: 603-609.

McKean, P. G., Keen, J. K., Smith, D. F. and Benson, F. E. (2001) Identification and characterisation of a RAD51 gene from *Leishmania major*, *Molecular and Biochemical Parasitology* **115**: 209-216.

McKerrow, J. H., Caffrey, C., Kelly, B., Loke, P. and Sajid, M. (2006) Proteases in parasitic diseases, *Annual Review of Pathology* **1**: 497-536.

McLuskey, K., Paterson, N. G., Bland, N. D., Isaacs, N. W. and Mottram, J. C. (2010) Crystal structure of *Leishmania major* oligopeptidase B gives insight into the enzymatic properties of a trypanosomatid virulence factor, *Journal of Biological Chemistry* **285**: 39249-39259.

Medina-Acosta, E. and Cross, G. A. (1993) Rapid isolation of DNA from trypanosomatid protozoa using a simple 'mini-prep' procedure, *Molecular and Biochemical Parasitology* **59**: 327-329.

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*, *Eukaryotic Cell* **7**: 684-697.

Militello, K. T., Refour, P., Comeaux, C. A. and Duraisingh, M. T. (2008) Antisense RNA and RNAi in protozoan parasites: working hard or hardly working?, *Molecular and Biochemical Parasitology* **157**: 117-126.

Miller, M. J., Wrightsman, R. A. and Manning, J. E. (1996) *Trypanosoma cruzi*:protective immunity in mice immunized with paraflagellar rod proteins is associated with a T-Helper type 1 response, *Experimental Parasitology* **84**: 156-167.

Mkunza, F., Olaho, W. M. and Powell, C. N. (1995) Partial protection against natural trypanosomiasis after vaccination with a flagellar pocket antigen from *Trypanosoma brucei rhodesiense*, *Vaccine* **13**: 151-154.

Montgomery, M. K., Xu, S. and Fire, A. (1998) RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*, *Proceedings of the National Academy of Sciences of the United States of America* **95**: 15502-15507.

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, *Molecular and Biochemical Parasitology* **157**: 160-168.

Mordmüller, B., Fendel, R., Kreidenweiss, A., Gille, C., Hurwitz, R., Metzger, W. G., Kun, J. F. J., Lamkemeyer, T., Nordheim, A. and Kremsner, P. G. (2006) Plasmodia express two threonine-peptidase complexes during asexual development, *Molecular and Biochemical Parasitology* **148**: 79-85.

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

Morty, R. E., Authié, E., Troeberg, L., Lonsdale-Eccles, J. D. and Coetzer, T. H. (1999a) Purification and characterisation of a trypsin-like serine oligopeptidase from *Trypanosoma congolense*, *Molecular and Biochemical Parasitology* **102**: 145-155.

Morty, R. E., Lonsdale-Eccles, J. D., Morehead, J., Caler, E. V., Mentele, R., Auerswald, E. A., Coetzer, T. H., Andrews, N. W. and Burleigh, B. A. (1999b) Oligopeptidase B from *Trypanosoma brucei*, a new member of an emerging subgroup of serine oligopeptidases, *Journal of Biological Chemistry* **274**: 26149-26156.

Morty, R. E., Troeberg, L., Powers, J. C., Ono, S., Lonsdale-Eccles, J. D. and Coetzer, T. H. T. (2000) Characterisation of the antitrypanosomal activity of peptidyl alpha-aminoalkyl phosphonate diphenyl esters, *Biochemical pharmacology* **60**: 1497-1504.

Morty, R. E., Lonsdale-Eccles, J. D., Mentele, R., Auerswald, E. A. and Coetzer, T. H. T. (2001) Trypanosome-derived oligopeptidase B is released into the plasma of infected rodents, where it persists and retains full catalytic activity, *Infection and Immunity* **69**: 2757-2761.

Morty, R. E. and Burleigh, B. A. (2004) Oligopeptidase B. In: *Handbook of Proteolytic Enzymes.* Barrett, A. J., Rawlings, N. D. and Woessner, J. F. (eds.), Elsevier, London. 1900-1905.

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

Morty, R. E., Shih, A. Y., Fülöp, V. and Andrews, N. W. (2005b) Identification of the reactive cysteine residues in oligopeptidase B from *Trypanosoma brucei*, *FEBS Letters* **579**: 2191-2196.

Morty, R. E., Vadasz, I., Bulau, P., Dive, V., Oliveira, V., Seeger, W. and Juliano, L. (2005c) Tropolysin, a new oligopeptidase from African trypanosomes, *Biochemistry* **44**: 14658-14669.

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, *Biochemical Journal* **394**: 635-645.

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 Letters* **258**: 211-215.

Mottram, J. C., Robertson, C. D., Coombs, G. H. and Barry, J. D. (1992) A developmentally regulated cysteine proteinase gene of *Leishmania mexicana*, *Molecular Microbiology* **6**: 1925-1932.

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, *Proceedings of the National Academy of Sciences of the United States of America* **93**: 6008-6013.

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, *Journal of Biological Chemistry* **272**: 14285-14293.

Mottram, J. C., Brooks, D. R. and Coombs, G. H. (1998) Roles of cysteine proteinases of trypanosomes and *Leishmania* in host-parasite interactions, *Current Opinion in Microbiology* **1**: 455-460.

Mottram, J. C., Coombs, G. H. and Alexander, J. (2004) Cysteine peptidases as virulence factors of *Leishmania*, *Current Opinion in Microbiology* **7**: 375-381.

Motyka, S. A. and Englund, P. T. (2004) RNA interference for analysis of gene function in trypanosomatids, *Current Opinion in Microbiology* **7**: 362-368.

Msangi, A. R., Saleh, K. M., Kiwia, N., Malele, I. I., Mussa, W. A., Mramba, F., Juma, K. G., Dyck, V. A., Vreysen, M. J. B., Parker, A. G., Feldmann, U., Zhu, Z. R. and Pan, H. (2000) Success in Zanzibar: eradication of tsetse. In: *Proceedings: Area-Wide Control of Fruit Flies and Other Insect Pests. International Conference on Area-Wide Control of Insect Pests, and the 5th International Symposium on Fruit Flies of Economic Importance.* Tan, K.-H. (eds.), Penerbit Universiti Sains Malaysia, Pulau Pinang. 57-66.

Munday, J. C., McLuskey, K., Brown, E., Coombs, G. H. and Mottram, J. C. (2011) Oligopeptidase B deficient mutants of *Leishmania major*, *Molecular and Biochemical Parasitology* **175**: 49-57.

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.

Mundodi, V., Kucknoor, A. S. and Gedamu, L. (2005) Role of *Leishmania (Leishmania) chagasi* amastigote cysteine protease in intracellular parasite survival: studies by gene disruption and antisense mRNA inhibition, *BioMed Central Molecular Biology* **6**: 3.

Murata, C. E. and Goldberg, D. E. (2003) *Plasmodium falciparum* Falcilysin, *Journal of Biological Chemistry* **278**: 38022-38028.

Murray, M., Morrison, W. I. and Whitelaw, D. D. (1982) Host susceptability to African trypanosomiasis: trypanotolerance, *Advances in Parasitology* **21**: 1-68.

Murray, M., Hirumi, H. and Moloo, S. K. (1985) Suppression of *Trypanosoma congolense*, *T. vivax* and *T. brucei* infection rates in tsetse flies maintained on goats immunized with uncoated forms of trypanosomes grown in vitro, *Parasitology* **91 (Pt 1)**: 53-66.

Murray, M., d'Ieteren, G. D. M. and Teale, A. J. (2004) Trypanotolerance. In: *The Trypanosomiases.* Maudlin, I., Holmes, P. H. and Miles, M. A. (eds.), CABI Publishing, Wallingford. 461-477.

Naessens, J., Teale, A. J. and Sileghem, M. (2002) Identification of mechanisms of natural resistance to African trypanosomiasis in cattle, *Veterinary Immunology and Immunopathology* **87**: 187-194.

Naessens, J., Leak, S. G., Kennedy, D. J., Kemp, S. J. and Teale, A. J. (2003) Responses of bovine chimaeras combining trypanosomosis resistant and susceptible genotypes to experimental infection with *Trypanosoma congolense*, *Veterinary Parasitology* **111**: 125-142.

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

Nagamune, K., Ohishi, K., Ashida, H., Hong, Y., Hino, J., Kangawa, K., Inoue, N., Maeda, Y. and Kinoshita, T. (2003) GPI Transamidase of *Trypanosoma brucei* Has Two Previously Uncharacterized (Trypanosomatid Transamidase 1 and 2) and Three Common Subunits, *Proceedings of the National Academy of Sciences of the United States of America* **100**: 10682-10687.

Namangala, B., De Baetselier, P., Noel, W., Brys, L. and Beschin, A. (2001) Alternative versus classical macrophage activation during experimental African trypanosomosis, *Journal of Leukocyte Biology* **69**: 387-396.

Ndung'u, J. M., Wright, N. G., Jennings, F. W. and Murray, M. (1992) Changes in atrial natriuretic factor and plasma renin activity in dogs infected with *Trypanosoma brucei*, *Parasitology Research* **78**: 553-556.

Ngo, H., Tschudi, C., Gull, K. and Ullu, E. (1998) Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*, *Proceedings of the National Academy of Sciences of the United States of America* **95**: 14687-14692.

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

Nikolskaia, O. V., Kim, Y. V., kovbasnjuk, O., Kim, K. J. and Grab, D. J. (2006b) Entry of *Trypanosoma brucei gambiense* into microvascular endothelial cells of the human blood-brain barrier, *International Journal of Parasitolgy* **36**: 513-519.

Noble, E. R. and Noble, G. A. (1982) Parasitology: the biology of animal parasites. In: Noble, E. R. and Noble, G. A. (eds.), Lea & Febiger, Philadelphia. 1- 522 pp.

Nok, A. J. and Balogun, E. O. (2003) A bloodstream *Trypanosoma congolense sialidase* could be involved in anemia during experimental trypanosomiasis, *Journal of Biochemistry* **133**: 725-730.

Nok, A. J., Nzelibe, H. C. and Yako, S. K. (2003) *Trypanosoma evansi* sialidase: surface localization, properties and hydrolysis of ghost red blood cells and brain cells-implications in trypanosomiasis, *Z Naturforsch C* **58**: 594-601.

Novina, C. D. and Sharp, P. A. (2004) The RNAi revolution, *Nature* **430**: 161-164.

Nwagwu, M., Okenu, D. M., Olusi, T. A. and Molokwu, R. I. (1988) *Trypanosoma brucei* releases proteases extracellularly, *Transactions of the Royal Society of Tropical Medicine and Hygiene* **82**: 577.

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

O'Donoghue, A. J., Mahon, C. S., Goetz, D. H., O'Malley, J. M., Gallagher, D. M., Zhou, M., Murray, P. G., Craik, C. S. and Tuohy, M. G. (2008) Inhibition of a secreted glutamic peptidase prevents growth of the fungus *Talaromyces emersonii*, *Journal of Biological Chemistry* **283**: 29186-29195.

Ojok, L., Kaeufer-Weiss, I. and Weiss, E. (2002) Distribution of *Trypanosoma congolense* in infected multimammate rats (*Mastomys coucha*): light and electron microscopical studies, *Veterinary Parasitology* **105**: 327-336.

Olenick, J. G., Wolff, R., Nauman, R. K. and McLaughlin, J. (1988) A flagellar pocket membrane fraction from *Trypanosoma brucei rhodesiense*: immunogold localization and nonvariant immunoprotection, *Infection and Immunity* **56**: 92-98.

Ommen, G., Lorenz, S. and Clos, J. (2009) One-step generation of double-allele gene replacement mutants in *Leishmania donovani*, *International Journal of Parasitolgy* **39**: 541-546.

Onishi, K., Li, Y., Ishii, K., Hisaeda, H., Tang, L., Duan, X., Dainichi, T., Maekawa, Y., Katunuma, N. and Himeno, K. (2004) Cathepsin L is crucial for a Th1-type immune response during Leishmania major infection, *Microbes and Infection* **6**: 468-474.

Opperdoes, F. R. and van Roy, J. (1982) The phospholipases of *Trypanosoma brucei* bloodstream forms and cultured procyclics, *Molecular and Biochemical Parasitology* **5**: 309-319.

Overath, P., Czichos, J. and Haas, C. (1986) The effect of citrate/cis-aconitate on oxidative metabolism during transformation of *Trypanosoma brucei*, *European Journal of Biochemistry* **160**: 175-182.

Overath, P., Ruoff, J., Stierhof, Y. D., Haag, J., Tichy, H., Dykova, I. and Lom, J. (1998) Cultivation of bloodstream forms of *Trypanosoma carassii*, a common parasite of freshwater fish, *Parasitology Research* **84**: 343-347.

Page, M. and Di Cera, E. (2008) Serine peptidases: Classification, structure and function, *Cellular and Molecular Life Sciences* **65**: 1220-1236.

Pandey, K. C., Singh, N., Arastu-Kapur, S., Bogyo, M. and Rosenthal, P. J. (2006) Falstatin, a cysteine protease inhibitor of *Plasmodium falciparum*, facilitates erythrocyte invasion, *PLoS Pathog* **2**: e117.

Panjaworayan, N. and Brown, C. M. (2011) Effects of HBV Genetic Variability on RNAi Strategies, *Hepatitis Research and Treatment* **2011**: 367908.

Parussini, F., García, M., Mucci, J., Agüero, F., Sánchez, D., Hellman, U., Åslund, L. and Cazzulo, J. J. (2003) Characterization of a lysosomal serine carboxypeptidase from *Trypanosoma cruzi*, *Molecular and Biochemical Parasitology* **131**: 11-23.

Pays, E., Vanhamme, L. and Perez-Morga, D. (2004) Antigenic variation in *Trypanosoma brucei*: facts, challenges and mysteries, *Current Opinion in Microbiology* **7**: 369-374.

Pays, E. (2006) The variant surface glycoprotein as a tool for adaptation in African trypanosomes, *Microbes and Infection* **8**: 930-937.

Peck, R. F., Shiflett, A. M., Schwartz, K. J., McCann, A., Hajduk, S. L. and Bangs, J. D. (2008) The LAMP-like protein p67 plays an essential role in the lysosome of African trypanosomes, *Molecular Microbiology* **68**: 933-946.

Pelle, R. and Murphy, N. B. (1993) Northern hybridization: rapid and simple electrophoretic conditions, *Nucleic Acids Research* **21**: 2783-2784.

Pillai, B., Cherney, M. M., Hiraga, K., Takada, K., Oda, K. and James, M. N. G. (2007) Crystal Structure of Scytalidoglutamic Peptidase with its First Potent Inhibitor Provides Insights into Substrate Specificity and Catalysis, *Journal of Molecular Biology* **365**: 343-361.

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

Pillay, D. (2011) Identification and characterisation of novel pathogenic factors of *Trypanosoma congolense*. School of Biochemistry, Genetics and Microbiology. University of KwaZulu-Natal, Pietermaritzburg.

Pollock, K. G., McNeil, K. S., Mottram, J. C., Lyons, R. E., Brewer, J. M., Scott, P., Coombs, G. H. and Alexander, J. (2003) The *Leishmania mexicana* cysteine protease, CPB2.8, induces potent Th2 responses, *Journal of Immunology* **170**: 1746-1753.

Proto, W. R., Castanys-Munoz, E., Black, A., Tetley, L., Moss, C. X., Juliano, L., Coombs, G. H. and Mottram, J. C. (2011) *Trypanosoma brucei* Metacaspase 4 Is a Pseudopeptidase and a Virulence Factor, *Journal of Biological Chemistry* **286**: 39914-39925.

Proudfoot, C. and McCulloch, R. (2005) Distinct roles for two RAD51-related genes in *Trypanosoma brucei* antigenic variation, *Nucleic Acids Res* **33**: 6906-6919.

Proudfoot, C. and McCulloch, R. (2006) *Trypanosoma brucei* DMC1 does not act in DNA recombination, repair or antigenic variation in bloodstream stage cells, *Molecular and Biochemical Parasitology* **145**: 245-253.

Qiu, S., Adema, C. M. and Lane, T. (2005) A computational study of off-target effects of RNA interference, *Nucleic Acids Research* **33**: 1834-1847.

Radwanska, M., Magez, S., Dumont, N., Pays, A., Nolan, D. and Pays, E. (2000a) Antibodies raised against the flagellar pocket fraction of *Trypanosoma brucei* preferentially recognize HSP60 in cDNA expression library, *Parasite Immunology* **22**: 639-650.

Radwanska, M., Magez, S., Michel, A., Stijlemans, B., Geuskens, M. and Pays, E. (2000b) Comparative Analysis of Antibody Responses against HSP60, Invariant Surface Glycoprotein 70, and Variant Surface Glycoprotein Reveals a Complex Antigen-Specific Pattern of Immunoglobulin Isotype Switching during Infection by *Trypanosoma brucei*, *Infection and Immunity* **68**: 848-860.

Radwanska, M., Guirnalda, P., De Trez, C., Ryffel, B., Black, S. and Magez, S. (2008) Trypanosomiasis-induced B cell apoptosis results in loss of protective anti-parasite antibody responses and abolishment of vaccine-induced memory responses, *PLoS Pathogens* **4**: e1000078.

Rawlings, N. D., Barrett, A. J. and Alan, J. B. (1994) Families of cysteine peptidases. In: *Methods in Enzymology.* Barrett, A. J., Rawlings, N. D. and Woessner, J. F. (eds.), Elsevier, London. 461-486.

Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. and Barrett, A. J. (2008) MEROPS: the peptidase database, *Nucleic Acids Research* **36**: 320-325.

Rawlings, N. D., Barrett, A. J. and Bateman, A. (2010) MEROPS: the peptidase database, *Nucleic Acids Res* **38**: D227-233.

Rawlings, N. D., Barrett, A. J. and Bateman, A. (2011) Asparagine Peptide Lyases, *Journal of Biological Chemistry* **286**: 38321-38328.

Rea, D. and Fülöp, V. (2006) Structure-function properties of prolyl oligopetidase family enzymes, *Cell Biochemistry and Biophysics* **44**: 349-365.

Reina-San-Martin, B., Degrave, W., Rougeot, C., Cosson, A., Chamond, N., Cordeiro-Da-Silva, A., Arala-Chaves, M., Coutinho, A. and Minoprio, P. (2000) A B-cell mitogen from a pathogenic trypanosome is a eukaryotic proline racemase, *Nature Medicine* **6**: 890-897.

Richardson, J. P., Beecroft, R. P., Tolson, D. L., Liu, M. K. and Pearson, T. W. (1988) Procyclin: an unusual immunodominant glycoprotein surface antigen from the procyclic stage of African trypanosomes, *Molecular and Biochemical Parasitology* **31**: 203-216.

Riviere, L., van Weelden, S. W., Glass, P., Vegh, P., Coustou, V., Biran, M., van Hellemond, J. J., Bringaud, F., Tielens, A. G. and Boshart, M. (2004) Acetyl:succinate CoA-transferase in procyclic *Trypanosoma brucei*. Gene identification and role in carbohydrate metabolism, *Journal of Biological Chemistry* **279**: 45337-45346.

Roberts, R. J., Bellfort, M., Bestor, T., Bhagwat, A. S., Bickle, T. A., Bitinaite, J., Blumenthal, R. M., Degtyarev, S. K., Dryden, D. T. F. and Dybvig, K. (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes, *Nucleic Acids Research* **31**: 1805.

Robinson, M. W., Dalton, J. P. and Donnelly, S. (2008) Helminth pathogen cathepsin proteases: it's a family affair, *Trends in Biochemical Sciences* **33**: 601-608.

Robinson, N. P., McCulloch, R., Conway, C., Browitt, A. and Barry, J. D. (2002) Inactivation of Mre11 Does Not Affect VSG Gene Duplication Mediated by Homologous Recombination in *Trypanosoma brucei*, *Journal of Biological Chemistry* **277**: 26185-26193.

Rogers, D. J. and Randolph, S. E. (2002) A response to the aim of eradicating tsetse from Africa, *Trends in parasitology* **18**: 534-536.

Ruepp, S., Furger, A., Kurath, U., Renggli, C. K., Hemphill, A., Brun, R. and Roditi, I. (1997) Survival of *Trypanosoma brucei* in the tsetse fly is enhanced by the expression of specific forms of procyclin, *Journal of Cell Biology* **137**: 1369-1379.

Rusconi, F., Durand-Dubief, M. and Bastin, P. (2005) Functional complementation of RNA interference mutants in trypanosomes, *BMC Biotechnology* **5**: 6.

Ryan, K. A. and Englund, P. T. (1989) Synthesis and processing of kinetoplast DNA minicircles in *Trypanosoma equiperdum*, *Molecular and Cellular Biology* **9**: 3212-3217.

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

Sanderson, S. J., Westrop, G. D., Scharfstein, J., Mottram, J. C. and Coombs, G. H. (2003) Functional conservation of a natural cysteine peptidase inhibitor in protozoan and bacterial pathogens, *FEBS Letters* **542**: 12-16.

Santana, J., Grellier, P., Schrevel, J. and Teixeira, A. (1997) A *Trypanosoma cruzi*-secreted 80 kDa proteinase with specificity for human collagen types I and IV, *Biochemical Journal* **325**: 129 - 137.

Santos, C. C., Coombs, G. H., Lima, A. P. and Mottram, J. C. (2007) Role of the *Trypanosoma brucei* natural cysteine peptidase inhibitor ICP in differentiation and virulence, *Molecular Microbiology* **66**: 991-1002.

Saravia, N. G., Escorcia, B., Osorio, Y., Valderrama, L., Brooks, D., Arteaga, L., Coombs, G., Mottram, J. and Travi, B. L. (2006) Pathogenicity and protective immunogenicity of cysteine proteinase-deficient mutants of *Leishmania mexicana* in non-murine models, *Vaccine* **24**: 4247-4259.

Scharfstein, J., Schmitz, V., Morandi, V., Capella, M. M. A., Lima, A. P. C. A., Morrot, A., Juliano, L. and Müller-Esterl, W. (2000) Host cell invasion by *Trypanosoma cruzi* is potentiated by activation of bradykinin B2 receptors, *The Journal of Experimental Medicine* **192**: 1289-1300.

Schliebs, W. (2006) Sleeping sickness: PEX and drugs, *Biochimica et Biophysica Acta* **1763**: 4-5.

Schmidt, A., Clayton, C. E. and Krauth-Siegel, R. L. (2002) Silencing of the thioredoxin gene in *Trypanosoma brucei brucei*, *Molecular and Biochemical Parasitology* **125**: 207-210.

Schofield, C. J. and Maudlin, I. (2001) Trypanosomiasis control, *International journal for parasitology* **31**: 614-619.

Schofield, L., Vivas, L., Hackett, F., Gerold, P., Schwarz, R. T. and Tachado, S. (1993) Neutralizing monoclonal antibodies to glycosylphosphatidylinositol, the dominant TNF-alpha-inducing toxin of *Plasmodium falciparum*: prospects for the immunotherapy of severe malaria, *Ann Trop Med Parasitol* **87**: 617-626.

Schofield, L. (2007) Rational approaches to developing an anti-disease vaccine against malaria, *Microbes and Infection* **9**: 784-791.

Scory, S., Caffrey, C. R., Stierhof, Y.-D., Ruppel, A. and Steverding, D. (1999) *Trypanosoma brucei*: Killing of Bloodstream Formsin Vitroandin Vivoby the Cysteine Proteinase Inhibitor Z-Phe-Ala-CHN2, *Experimental Parasitology* **91**: 327-333.

Sharp, P. A. (2001) RNA interference, *Genes & Development* **15**: 485-490.

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.

Shi, H., Ullu, E. and Tschudi, C. (2004a) Function of the Trypanosome Argonaute 1 protein in RNA interference requires the N-terminal RGG domain and arginine 735 in the Piwi domain, *Journal of Biological Chemistry* **279**: 49889-49893.

Shi, H., Djikeng, A., Tschudi, C. and Ullu, E. (2004b) Argonaute protein in the early divergent eukaryote *Trypanosoma brucei*: control of small interfering RNA accumulation and retroposon transcript abundance, *Molecular and Cellular Biology* **24**: 420-427.

Shi, H., Tschudi, C. and Ullu, E. (2006a) An unusual Dicer-like1 protein fuels the RNA interference pathway in *Trypanosoma brucei*, *RNA* **12**: 2063-2072.

Shi, H., Tschudi, C. and Ullu, E. (2006b) Functional replacement of *Trypanosoma brucei* Argonaute by the human slicer Argonaute2, *RNA* **12**: 943-947.

Sijwali, P. S., Shenai, B. R., Gut, J., Singh, A. and Rosenthal, P. J. (2001) Expression and characterization of the *Plasmodium falciparum* haemoglobinase falcipain-3, *Biochemical Journal* **360**: 481-489.

Sijwali, P. S. and Rosenthal, P. J. (2004) Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by *Plasmodium falciparum*, *Proceedings of the National Academy of Sciences of the United States of America* **101**: 4384-4389.

Silva-Lopez, R. E., Morgado-Diaz, J. A., Chavez, M. A. and Giovanni-De-Simone, S. (2007) Effects of serine protease inhibitors on viability and morphology of *Leishmania (Leishmania) amazonensis* promastigotes, *Parasitology Research* **101**: 1627-1635.

Silva, A. M., Lee, A. Y., Gulnik, S. V., Maier, P., Collins, J., Bhat, T. N., Collins, P. J., Cachau, R. E., Luker, K. E., Gluzman, I. Y., Francis, S. E., Oksman, A., Goldberg, D. E. and Erickson, J. W. (1996) Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from *Plasmodium falciparum*, *Proceedings of the National Academy of Sciences of the United States of America* **93**: 10034-10039.

Simpson, L. (1987) The mitochondrial genome of kinetoplastid protozoa: genomic organisation, transcription, replication, and evolution, *Annual Review of Microbiology* **41**: 363-382.

Skelly, P. J., Da'dara, A. and Harn, D. A. (2003) Suppression of cathepsin B expression in *Schistosoma mansoni* by RNA interference, *International journal for parasitology* **33**: 363-369.

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 in parasitology*: In press.

Solberg, H., Rinkenberger, J., Dano, K., Werb, Z. and Lund, L. R. (2003) A functional overlap of plasminogen and MMPs regulates vascularization during placental development, *Development* **130**: 4439-4450.

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 Letters* **311**: 124-127.

Stechmiller, J., Cowan, L. and Schultz, G. (2010) The Role of Doxycycline as a Matrix Metalloproteinase Inhibitor for the Treatment of Chronic Wounds, *Biological Research For Nursing* **11**: 336-344.

Stijlemans, B., Guilliams, G., Raes, G., Beschin, A., Magez, S. and De Baetselier, P. (2007a) African trypanosomosis: from immune escape and immunopathology to immune intervention, *Veterinary Parasitology* **148**: 3-13.

Stijlemans, B. t., Baral, T. N., Guilliams, M., Brys, L., Korf, J., Drennan, M., Van Den Abbeele, J., De Baetselier, P. and Magez, S. (2007b) A glycosylphosphatidylinositol-based treatment alleviates trypanosomiasis-associated immunopathology, *Journal of Immunology* **179**: 4003-4014.

Sturtevant, J. E. and Balber, A. E. (1983) Externally disposed membrane polypeptides of intact and protease-treated *Trypanosoma lewisi* correlated with sensitivity to alternate complement pathway-mediated lysis, *Infection and Immunity* **42**: 869-875.

Sui, G., Soohoo, C., Affar el, B., Gay, F., Shi, Y. and Forrester, W. C. (2002) A DNA vector-based RNAi technology to suppress gene expression in mammalian cells, *Proceedings of the National Academy of Sciences of the United States of America* **99**: 5515-5520.

Sung, P. and Klein, H. (2006) Mechanism of homologous recombination: mediators and helicases take on regulatory functions, *Nature Reviews Molecular Cell Biology* **7**: 739-750.

Swallow, B. M. (2000) Impacts of trypanosomiasis on African agriculture. In: *PAAT Technical and Scientific Series.* FAO/WHO/IAEA/OAU-IBAR (eds.), Food and Agriculture Organization of the United Nations, Rome. 1-52.

Swenerton, R. K., Zhang, S., Sajid, M., Medzihradszky, K. F., Craik, C. S., Kelly, B. L. and McKerrow, J. H. (2011) The oligopeptidase B of *Leishmania* regulates parasite enolase and immune evasion, *Journal of Biological Chemistry* **286**: 429-440.

Tabara, H., Grishok, A. and Mello, C. C. (1998) RNAi in *C. elegans*: soaking in the genome sequence., *Science* **282**: 430-431.

Tabel, H., Kaushik, R. S. and Uzonna, J. E. (2000) Susceptibility and resistance to *Trypanosoma congolense* infections, *Microbes and Infection* **2**: 1619-1629.

Tan, K. S. W., Leal, S. T. G. and Cross, G. A. M. (2002) *Trypanosoma brucei* MRE11 is non-essential but influences growth, homologous recombination and DNA double-strand break repair, *Molecular and Biochemical Parasitology* **125**: 11-21.

Tartof, K. D. and Hobbs, C. A. (1987) Improved media for growing plasmid and cosmid clones, *Bethesda Research laboratories* **Focus 9**: 12.

Taylor, J. E. and Rudenko, G. (2006) Switching trypanosome coats: what's in the wardrobe?, *Trends in Genetics* **22**: 614-620.

Taylor, K. A., Lutje, V., Kennedy, D., Authié, E., Boulangé, A., Logan-Henfrey, L., Gichuki, B. and Gettinby, G. (1996) *Trypanosoma congolense*: B-lymphocyte responses differ between trypanotolerant and trypanosusceptible cattle., *Experimental Parasitology* **83**: 106-116.

Taylor, K. A. (1998) Immune responses of cattle to African Trypanosomes: protective or pathogenic, *International journal for parasitology* **28**: 219-240.

Taylor, K. A. and Mertens, B. (1999) Immune response of cattle infected with African trypanosomes, *Memórias do Instituto Oswaldo Cruz* **94**: 239-244.

Taylor, K. A. and Authié, E. M.-L. (2004) Pathogenesis of Animal Trypanosomiasis. In: *The Trypanosomiases.* Maudlin, I., Holmes, P. H. and Miles, M. A. (eds.), CABI Publishing, Wallingford. 331-353.

Tazeh, N. N., Silverman, J. S., Schwartz, K. J., Sevova, E. S., Sutterwala, S. S. and Bangs, J. D. (2009) Role of AP-1 in developmentally regulated lysosomal trafficking in *Trypanosoma brucei*, *Eukaryot Cell* **8**: 1352-1361.

Teixeira, A. R. L. and Santana, J. M. (1989) *Trypanosoma Cruzi*: Endocytosis and Degradation of Specific Antibodies by Parasite Forms, *The American Journal of Tropical Medicine and Hygiene* **40**: 165-170.

ten Asbroek, A. L. M. A., Ouellette, M. and Borst, P. (1990) Targeted insertion of the neomycin phosphotransferase gene into the tubulin gene cluster of *Trypanosoma brucei*, *Nature* **348**: 174-175.

Tetaert, D., Soudan, B., Huet-Duvillier, G., Degand, P. and Boersma, A. (1993) Unusual cleavage of peptidic hormones generated by trypanosome enzymes released in infested rat serum, *International Journal of Peptide and Protein Research* **41**: 147-152.

Tetley, L., Turner, C. M., Barry, J. D., Crowe, J. S. and Vickerman, K. (1987) Onset of expression of the variant surface glycoproteins of *Trypanosoma brucei* in the tsetse fly studied using immunoelectron microscopy, *Journal of Cell Science* **87 (Pt 2)**: 363-372.

Tizard, I., Nielsen, K. H., Seed, J. R. and Hall, J. E. (1978) Biologically active products from African Trypanosomes, *Microbiology Reviews* **42**: 664-681.

Tosomba, O. M., Coetzer, T. H. T. and Lonsdale-Eccles, J. D. (1996) Localisation of acid phosphatase activity on the surface of bloodstream forms of *Trypanosoma congolense*, *Experimental Parasitology* **84**: 429-438.

Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from from polyacrylamide gels to nitrocellulose sheets: procedure and some applications, *Proceedings of the National Academy of Sciences of the United States of America* **76**: 4350-4354.

Tran, T., Büscher, P., Vandenbussche, G., Wyns, L., Messens, J. and De Greve, H. (2008) Heterologous expression, purification and characterisation of the extracellular domain of trypanosome invariant surface glycoprotein ISG75, *Journal of Biotechnology* **135**: 247-254.

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

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

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*, *Experimental Parasitology* **91**: 349-355.

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*, *Molecular Medicine* **6**: 660-669.

Tsuji, A., Yuasa, K. and Matsuda, Y. (2004) Identification of oligopeptidase B in higher plants. Purification and characterization of oligopeptidase B from quiescent wheat embryo, *Triticum aestivum*, *Journal of Biochemistry* **136**: 673-681.

Tsuji, A., Yoshimoto, T., Yuasa, K. and Matsuda, Y. (2006) Protamine: a unique and potent inhibitor of oligopeptidase B, *Journal of Peptide Research* **12**: 65-71.

Turk, V., Stoka, V. and Turk, D. (2008) Cystatins: biochemical and structural properties, and medical relevance, *Frontiers in Bioscience* **13**: 5406-5420.

Ullu, E., Djikeng, A., Shi, H. and Tschudi, C. (2002) RNA interference: advances and questions, *Philosophical Transactions of the Royal Society B: Biological Sciences* **357**: 65-70.

Ullu, E., Tschudi, C. and Chakraborty, T. (2004) RNA interference in protozoan parasites, *Cellular Microbiology* **6**: 509-519.

Van Den Berghe, L. and Zaghi, A. J. (1963) Wild Pigs as Hosts of *Glossina vanhoofi* Henrard and *Trypanosoma suis* Ochmann in the Central African Forest, *Nature* **197**: 1126-1127.

Velez, I. D., Gilchrist, K., Arbelaez, M. P., Rojas, C. A., Puerta, J. A., Antunes, C. M., Zicker, F. and Modabber, F. (2005) Failure of a killed *Leishmania amazonensis* vaccine against American cutaneous leishmaniasis in Colombia, *Transactions of the Royal Society of Tropical Medicine and Hygiene* **99**: 593-598.

Vickerman, K. (1982) Zoomastigophora. In: *Synopsis and Classification of Living Organisms.* Parker, S. P. (eds.), McGraw-Hill, New York. 496-508.

Vickerman, K. (1985) Developmental cycles and biology of pathogenic trypanosomes, *British Medical Bulletin* **41**: 105-114.

Vickerman, K., Myler, P. J. and Stuart, D. K. (1993) African trypanosomosis. In: *Immunology and molecular biology of parasitic infections.* Warren, K. S. (eds.), Blackwell scientific publications, 170-212.

Vickerman, K. (2000) Order *Kinetoplastea Honigberg*, 1963. In: *An Illustrated Guide to the Protozoa.* Lee, J. J., Leedale, G.F. & Bradbury, P. (eds.), Society of Protozoologists, Lawrence, Kansas. 1159-1185.

Voet, D., Voet, J. G. and Pratt, C. (2006) *Fundamentals of biochemistry : life at the molecular level*: John Wiley & Sons 14-23.

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, *Journal of Economic Entomology* **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, *Biochemical Journal* **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, *Journal of Biological Chemistry* **275**: 40174-40179.

Weinberg, M. S. and Arbuthnot, P. (2010) Progress in the use of RNA interference as a therapy for chronic hepatitis B virus infection, *Genome Med* **2**: 28.

Wilkinson, S. R., Prathalingam, S. R., Taylor, M. C., Ahmed, A., Horn, D. and Kelly, J. M. (2006) Functional characterisation of the iron superoxide dismutase gene repertoire in *Trypanosoma brucei*, *Free radical biology and medicine* **40**: 198-209.

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*, *Molecular and Biochemical Parasitology* **99**: 89-101.

Wrightsman, R. A., Miller, M. J., Saborio, J. L. and Manning, J. E. (1995) Pure paraflagellar rod protein protects mice against *Trypanosoma cruzi* infection, *Infection and Immunity* **63**: 122-125.

Wurst, M., Robles, A., Po, J., Luu, V.-D., Brems, S., Marentije, M., Stoitsova, S., Quijada, L., Hoheisel, J., Stewart, M., Hartmann, C. and Clayton, C. (2009) An RNAi screen of the RRM-domain proteins of *Trypanosoma brucei*, *Molecular and Biochemical Parasitology* **163**: 61-65.

Xu, D., Brandan, C. P., Basombrio, M. A. and Tarleton, R. L. (2009) Evaluation of high efficiency gene knockout strategies for *Trypanosoma cruzi*, *BMC Microbiology* **9**: 90.

Yoshida, H., Okamoto, K., Iwamoto, T., Sakai, E., Kanaoka, K., Hu, J.-P., Shibata, M., Hotokezaka, H., Nishishita, K., Mizuno, A. and Kato, Y. (2006) Pepstatin A, an aspartic proteinase inhibitor, suppresses RANKL-induced osteoclast differentiation, *Journal of Biochemistry* **139**: 583-590.

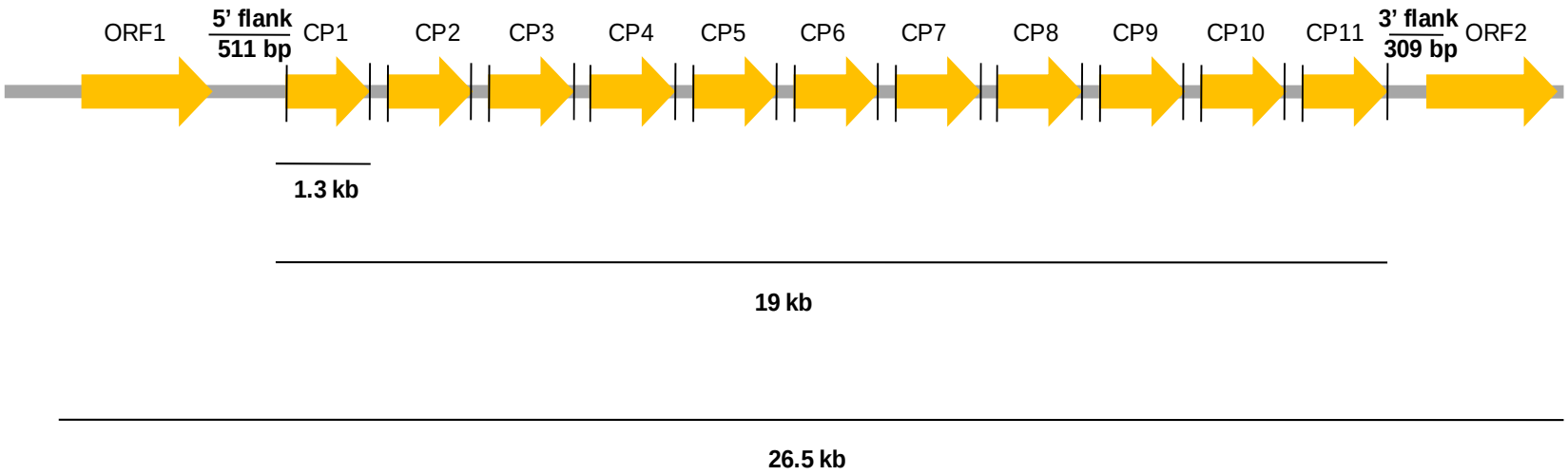
Zambrano-Villa, S., Rosales-Borjas, D., Carrero, J. C. and Ortiz-Ortiz, L. (2002) How protozoan parasites evade the immune response, *Trends in parasitology* **18**: 272-278.

Zhang, T., Maekawa, Y., Sakai, T., Nakano, Y., Ishii, K., Hisaeda, H., Dainichi, T., Asao, T., Katunuma, N. and Himeno, K. (2001) Treatment with cathepsin L inhibitor potentiates Th2-type immune response in *Leishmania major*-infected BALB/c mice, *International Immunology* **13**: 975-982.

APPENDIX

Appendix 1: Arrangement of *Tb*CATL genes within the *Tb*CATL locus in *T. brucei* 427 Lister strain genome

(GeneDB *http://www.genedb.org/genedb/tbrucei*)



Appendix 2: *T. congolense Tco*CATL sequence and 5' region flanking *Tco*CATL sequence alignment within contigs 0001218 and 0002521

(GeneDB *http://www.genedb.org/genedb/tcongolense*)

```
contig 0001218   AAGTCGCAAC AGTTGCACGC GGGTGAAGGC CACACAACAG AACAAGAGAG ATCGCTTTCC ACATCCTTCA 7490
5' overhang      ---------- -----CACGC GGGTGAAGGC CACACAACAG AACAAGAGAG ATCGCTTTCC ACATCCTTCA 55
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   ACCAAGCTGA AGAAGAACTG GAATGCGTCA TTGAGTGCTC GCTTCATCTA CAACTTCAAG CCTTACTATT 7560
5' overhang      ACCAAGCTGA AGAAGAACTG GAATGCGTCA TTGAGTGCTC GCTTCATCTA CAACTTCAAG CCTTACTATT 125
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   ATGCACCACA TTGCCCCAAG ACATGGTCCA CGGGTTTGAC ACATGGTACG TGGAAGCAAC ATCAAACATC 7630
5' overhang      ATGCACCACA TTGCCCCAAG ACATGGTCCA CGGGTTTGAC ACATGGTACG TGGAAGCAAC ATCAAACATC 195
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   GTTGATTTAG CACTTGGTGC ACCATGGCCA ATAGTGCGTC ACCTTCGGGC CGTTGGCTTC GCCTTCTGCC 7700
5' overhang      GTTGATTTAG CACTTGGTGC ACCATGGCCA ATAGTGCGTC ACCTTCGGGC CGTTGGCTTC GCCTTCTGCC 265
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   TGATTCCACC ACACCAACAG CATCAATCAC TCCTCATTGC TGTTGGTCTA TTTGCCACAC TAACACAACG 7770
5' overhang      TGATTCCACC ACACCAACAG CATCAATCAC TCCTCATTGC TGTTGGTCTA TTTGCCACAC TAACACAACG 335
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
```

```
contig 0001218   CCAGGAAGAG AAAAACAAGT CACACAAAGC TGGCGCCACA GAAACTAAAG GAGCACTTAA AACCAACATT 7840
5' overhang      CCAGGAAGAG AAAAACAAGT CACACAAAGC TGGCGCCACA GAAACTAAAG GAGCACTTAA AACCAACATT 405
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
contig 0001218   ATTATAGGCG CAGGATGGTG GGAGATCTTC AAAATATGTG TGCTGCACGA TTCCATTATT CCAGCACAAC 7910
5' overhang      ATTATAGGCG CAGGATGGTG GGAGATCTTC AAAATATGTG TGCTGCACGA TTCCATTATT CCAGCACAAC 475
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   TGCATAACGA TATATGGCAC CTAGTCGAGC GAAGCGAACG AGCACTTAGA ACAAGTCACT CGATACGCCT 7980
5' overhang      TGCATAACGA TATATGGCAC CTAGTCGAGC GAAGCGAACG AGCACTTAGA ACAAGTCACT CGATACGCCT 545
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   ACAAAACACA CAAAATAAAA ACAGCAGCAA GGACCACATT GACACACAAC CGAATCAAAA CAAAAACAAC 8050
5' overhang      ACAAAACACA CAAAATAAAA ACAGCAGCAA GGACCACATT GACACACAAC CGAATCAAAA CAAAAACAAC 615
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   CAAATGAAAC CAAGATTACT GTCACTAACA AGTAGAATAA TACACACAAC TAAACACTTA ATAAATAAAA 8120
5' overhang      CAAATGAAAC CAAGATTACT GTCACTAACA AGTAGAATAA TACCCACAAC TAAACACTTA ATAAATAAAA 685
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   CCAGGGAACA ATGGAACGGA AAAAAAGAAG GAAAACAAAA CCCTTTCACA GCAAGGAAAC ATCACACACC 8190
5' overhang      CCAGGGAACA ATGGAACGGA AAAAAAGAAG GAAAACAAAA CCCTTTCACA GCAAGGAAAC ATCACACACC 755
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
```

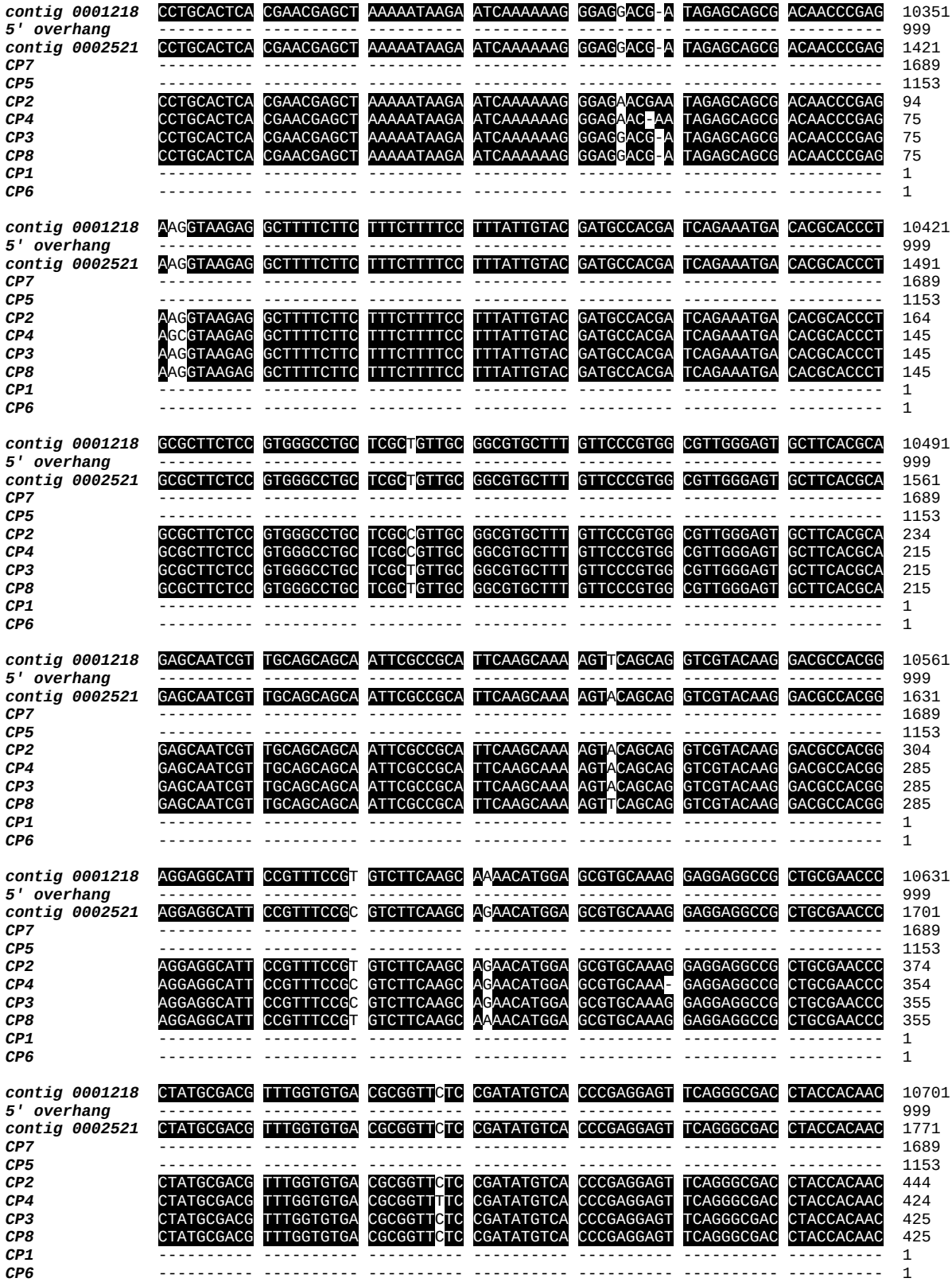
```
contig 0001218   CTATAAAATA ATCAGCAAAA AAAACAAACC AACATCATCT CTACAAAAGA AAAAAATAGC ACTATGGTGG 8260
5' overhang      CTATAAAATA ATCAGCAAAA AAAACAAACC AACATCATCT CTACAAAAGA AAAAAATAGC ACTATGGTGG 825
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   AAAATAGAAA TTAACCGACC ACAATAAAAG AAACAAATAA AAAACAAAGA ACAGCAACAG TAAC-AAAAA 8329
5' overhang      AAAATAGAAA TTAACCGACC ACCATAAAAG AATCCAATAA AAAACAAACC ACAGCCACAG TAACAAAAAA 895
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   AACATTACTT ACATTCTAAC ACCAGCAACC ATCTCTCCCT GTATGTCACA CCGTACGTTG GCGCAGTGTA 8399
5' overhang      AACATGACTT ACATTCTAAC ACCAGCAACA ATCTCTCCCT GTATGTCACA CCGTACGTTG GCGCAGTGTA 965
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              ---------- ---------- ---------- ---------- ---------- ---------- -----GTGTA 5
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   TCCTGCACTC ACGAACGAGC TAAAAATAAG AATCAAAAAA GGGAGAACGA TAGAGCAGCG ACAACCCGAG 8469
5' overhang      TCCTGCACTC ACGAACGAGC TAAAAATAAG AATA------ ---------- ---------- ---------- 999
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              TCCTGCACTC ACGAACGAGC TAAAAATAAG AATCAAAAAA GGGAGAACGA TAGAGCAGCG ACAACCCGAG 75
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   AAGGTAAGAG GCTTTTCTTC TTTCTTTTCC TTTATTGTAC GATGCCACGA TCAGAAATGA CACGCACCCT 8539
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              AAGGTAAGAG GCTTTTCTTC TTTCTTTTCC TTTATTGTAC GATGCCACGA TCAGAAATGA CACGCACCCT 145
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   GCGCTTCTCC GTGGGCCTGC TCGCCGTTGC GGCGTGCTTT GTTCCCGTGG CGTTGGGAGT GCTTCACGCA 8609
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7              GCGCTTCTCC GTGGGCCTGC TCGCCGTTGC GGCGTGCTTT GTTCCCGTGG CGTTGGGAGT GCTTCACGCA 215
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
```

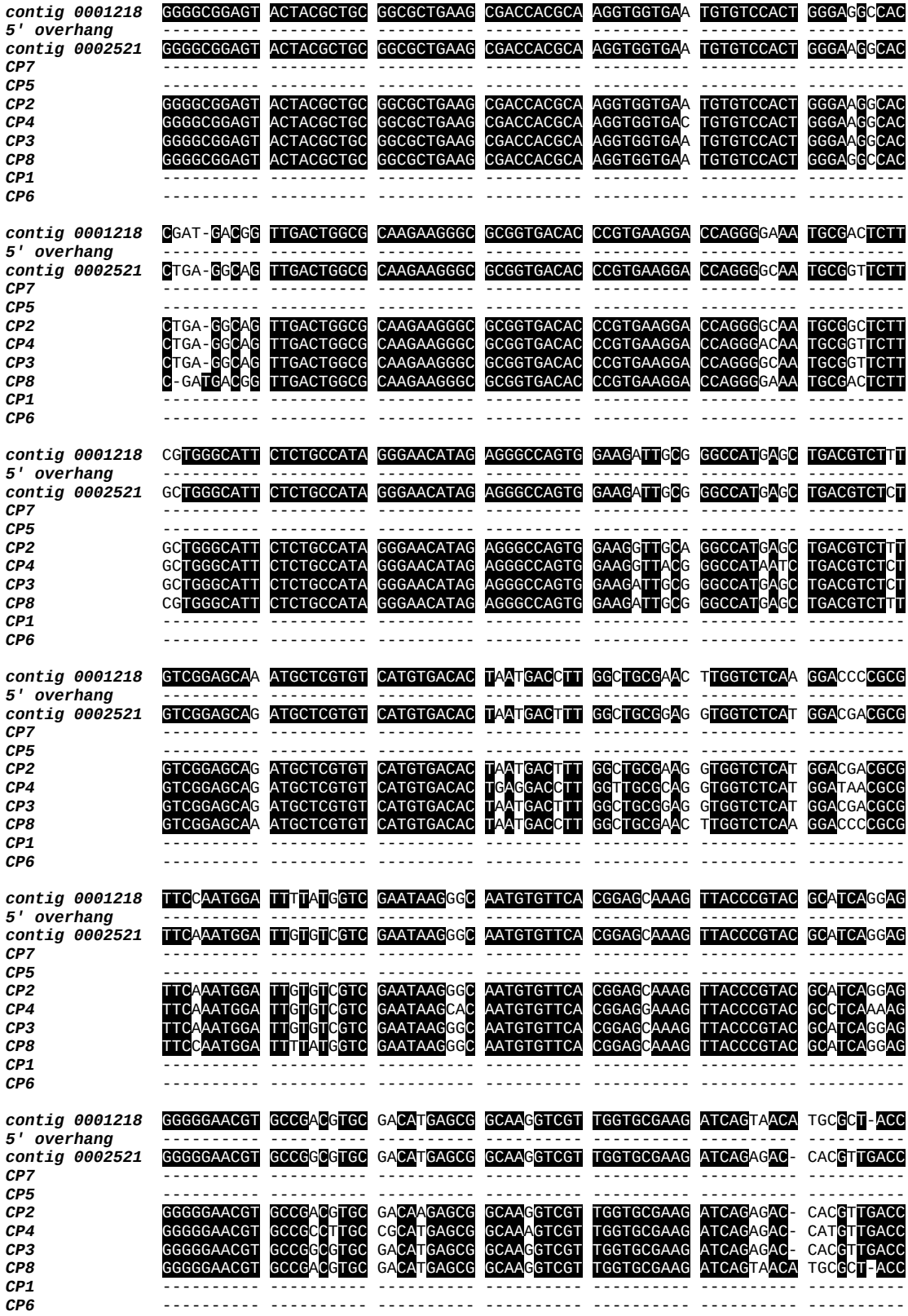
```
contig 0001218  GAGCAATCGT TGCAGCAGCA ATTCGCCGCA TTCAAGCAAA AGTACAGCAG GTCGTACAAG GACGCCACGG 8679
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7             GAGCAATCGT TGCAGCAGCA ATTCGCCGCA TTCAAGCAAA AGTACAGCAG GTCGTACAAG GACGCCACGG 285
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  AGGAGGCATT CCGTTTCCGC GTCTTCAAGC AGAACATGGA GCGTGCAAAG GAGGAGGCCG CTGCGAACCC 8749
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7             AGGAGGCATT CCGTTTCCGC GTCTTCAAGC AGAACATGGA GCGTGCAAAG GAGGAGGCCG CTGCGAACCC 355
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  CTATGCGACG TTTGGTGTGA CGCGGTTCTC CGATATGTCA CCCGAGGAGT TCAGGGCGAC CTACCACAAC 8819
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7             CTATGCGACG TTTGGTGTGA CGCGGTTCTC CGATATGTCA CCCGAGGAGT TCAGGGCGAC CTACCACAAC 425
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  GGGGCGGAGT ACTACGCTGC GGCGCTGAAG CGACCACGCA AGGTGGTGAA TGTGTCCACT GGGAAGGCAC 8889
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP7             GGGGCGGAGT ACTACGCTGC GGCGCTGAAG CGACCACGCA AGGTGGTGAA TGTGTCCACT GGGAAGGCAC 495
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  CTGAGGCAGT TGACTGGCGC AAGAAGGGCG CGGTGACACC CGTGAAGGAC CAGGGGAAAT GCGACTCTTC 8959
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  ---------- ---------- ---------- ----------     -----CGGAC CAGGGGCAAT GCGACTCTTC 25
CP7             CTGAGGCAGT TGACTGGCGC AAGAAGGGCG CGGTGACACC CGTGAAGGAC CAGGGGAAAT GCGACTCTTC 565
CP5             ---------- ---------- ---------- ----------       ----GGAC CAGGGGCAAT GCGACTCTTC 24
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  GTGGGCATTC ACTGTCATAG GGAACATAGA GGGCCAGTGG AAGATTGCGG GCCATGAGCT GACGTCTCTG 9029
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  GTGGGCATTC ACTGTCACGG GGAACATAGA GGGCCAGTGG AAGATTGCGG GCCATGAGCT GACGTCTCTG 95
CP7             GTGGGCATTC ACTGTCATAG GGAACATAGA GGGCCAGTGG AAGATTGCGG GCCATGAGCT GACGTCTCTG 635
CP5             GTGGGCATTC ACTGTCACGG GGAACATAGA GGGCCAGTGG AAGATTGCGG GCCATGAGCT GACGTCTCTG 94
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
```

```
contig 0001218   TCGGAGCAGA TGCTCGTGTC ATGTGACACT AATGACCTTG GCTGCCGAGC TGGCTTCATG GACACCGCGT 9099
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   TCGGAGCAGA TGCTCGTGTC ATGTGACACT AATGACCTTG GCTGCCGAGC TGGCTTCATG GACACCGCGT 165
CP7              TCGGAGCAGA TGCTCGTGTC ATGTGACACT AATGACCTTG GCTGCCGAGC TGGCTTCATG GACACCGCGT 705
CP5              TCGGAGCAGA TGCTCGTGTC ATGTGACACT AATGACCTTG GCTGCCGAGC TGGCTTCATG GACACCGCGT 164
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   TCAAATGGAT TGTGTCGTCG AATAATGGCA ATGTGTTCAC GGAGCAAAGT TACCCGTACG CATCAGGAGG 9169
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   TCAAATGGAT TGTGTCGTCG AATAATGGCA ATGTGTTCAC GGAGCAAAGT TACCCGTACG CATCAGGAGG 235
CP7              TCAAATGGAT TGTGTCGTCG AATAATGGCA ATGTGTTCAC GGAGCAAAGT TACCCGTACG CATCAGGAGG 775
CP5              TCAAATGGAT TGTGTCGTCG AATAATGGCA ATGTGTTCAC GGAGCAAAGT TACCCGTACG CATCAGGAGG 234
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   GGGGAACGTG CCGACGTGCA ACAAGAGCGG CAAGGTCGTT GGTGCGAACA TCGACGACCA CGTTCACATT 9239
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   GGGGAACGTG CCGACGTGCA ACAAGAGCGG CAAGGTCGTT GGTGCGAACA TCGACGACCA CGTTCACATT 305
CP7              GGGGAACGTG CCGACGTGCA ACAAGAGCGG CAAGGTCGTT GGTGCGAACA TCGACGACCA CGTTCACATT 845
CP5              GGGGAACGTG CCGACGTGCA ACAAGAGCGG CAAGGTCGTT GGTGCGAACA TCGACGACCA CGTTCACATT 304
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   CCGGAGGACG AAAACGCAAT AGCCGAGTGG CTGGCAAAAA AAGGGCCCGT CGCGATTGCC GTGGACGCCA 9309
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   CTAGACAACG AAAACGCAAT AGCCGAGTGG CTGGCAAAGA AAGGGCCCGT CGCGATTGCC GTTGACGCCA 375
CP7              CCGGAGGACG AAAACGCAAT AGCCGAGTGG CTGGCAAAAA AAGGGCCCGT CGCGATTGCC GTGGACGCCA 915
CP5              CTAGACAACG AAAACGCAAT AGCCGAGTGG CTGGCAAAGA AAGGGCCCGT CGCGATTGCC GTTGACGCCA 374
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   CCAGTTTCCA GAGCTACACG GGCGGAGTGT TGACATCATG CATCTCTAAG GAAGTGAACT CTGCTGCGCT 9379
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   CCAGTTTCCA GAGCTACAC- GGCGGAGTGT TGACATCATG CATCTCCGAG CATCTGGACC ATGGTGTGCT 444
CP7              CCAGTTTCCA GAGCTACACG GGCGGAGTGT TGACATCATG CATCTCTAAG GAAGTGAACT CTGCTGCGCT 985
CP5              CCAGTTTCCA GAGCTACAC- GGCGGAGTGT TGACATCATG CATCTCCGAG CATCTGGACC ATGGTGTGCT 443
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   CCTCGTGGGG TATGACGACA CAAGCAAGCC ACCATACTGG ATTATCAAGA ACTCATGGAG CAAGGGATGG 9449
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   CCTCGTGGGG TATGACGACA CAAGCAAGCC ACCATACTGG ATTATCAAGA ACTCATGGAG CAAGGGATGG 514
CP7              CCTCGTGGGG TATGACGACA CAAGCAAGCC ACCATACTGG ATTATCAAGA ACTCATGGAG CAAGGGATGG 1055
CP5              CCTCGTGGGG TATGACGACA CAAGCAAGCC ACCATACTGG ATTATCAAGA ACTCATGGAG CAAGGGATGG 513
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
```

```
contig 0001218   GGCGAGGAGG GTTACATCCG CATTGAGAAG GGCACAAACC AATGTCGTAT GAAAGAATAC GTGAGCTCCG 9519
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   GGCGAGGAGG GTTACATCCG CATTGAGAAG GGCACAAACC AATGTCTTAT GAAAAATCTC CCAAGCTCCG 584
CP7              GGCGAGGAGG GTTACATCCG CATTGAGAAG GGCACAAACC AATGTCGTAT GAAAGAATAC GTGAGCTCCG 1125
CP5              GGCGAGGAGG GTTACATCCG CATTGAGAAG GGCACAAACC AATGTCTTAT GAAAAATCTC CCAAGCTCCG 583
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   CCGTTGTCAG CGGCCCTCCG CCTCCG---C ---CCCCGAC TCCGACCTTC ACGCAGGAGT TGTGCGAGGG 9583
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   CCGTTGTCAG CGTCCCTCCG CCTCCGCCTC CACCCCCGAC TCCGACCTTC ACGCAGTCGT TCTGCCCCAA 654
CP7              CCGTTGTCAG CGGCCCTCCG CCTCCG---C ---CCCCGAC TCCGACCTTC ACGCAGGAGT TGTGCGAGGG 1189
CP5              CCGTTGTCAG CGTCCCTCCG CCTCCGCCTC CACCCCCGAC TCCGACCTTC ACGCAGTCGT TCTGCCCCAA 653
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   TGCCGAATGC CAGAGTAAGT GCACCAAAGC CACATTCCCA ACGGGCAAGT GCGTGCAGCT CAGCGGCGCC 9653
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   CGCAAAATGC CAGAGTAGGT GCACCAAAGC CACATTCCCA ACGGGCAAGT GCGTGCAGCT CAGCGACACC 724
CP7              TGCCGAATGC CAGAGTAAGT GCACCAAAGC CACATTCCCA ACGGGCAAGT GCGTGCAGCT CAGCGGCGCC 1259
CP5              CGCAAAATGC CAGAGTAGGT GCACCAAAGC CACATTCCCA ACGGGCAAGT GCGTGCAGCT CAGCGACACC 723
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   GGTTCTGTCA TCGCCTCGTG CGGCTCCAAT AACCTCACGC AGATCGTCTA CCCGTTGAGC AGCTCCTGCA 9723
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   GGCTCTGCCA TCGCCTCGTG CGGCTCCAAT AACCTCACGC AGATCGTCTA CCCGTTGAGC AGCTCCTGCA 794
CP7              GGTTCTGTCA TCGCCTCGTG CGGCTCCAAT AACCTCACGC AGATCGTCTA CCCGTTGAGC AGCTCCTGCA 1329
CP5              GGCTCTGCCA TCGCCTCGTG CGGCTCCAAT AACCTCACGC AGATCGTCTA CCCGTTGAGC AGCTCCTGCA 793
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   GCGGCTTCTC CGTTCCGTTG ACTGTGCCAC TGGACAAGTG CCTGCCCATT GTGATTGGAT CCGTGGCTTA 9793
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   GCGGCTACTC CGTTCCGTTG ACTGTGCCAC TGGACAAGTG CCTGCCCATT GTGATTGGAT CCATGATGTA 864
CP7              GCGGCTTCTC CGTTCCGTTG ACTGTGCCAC TGGACAAGTG CCTGCCCATT GTGATTGGAT CCGTGGCTTA 1399
CP5              GCGGCTACTC CGTTCCGTTG ACTGTGCCAC TGGACAAGTG CCTGCCCATT GTGATTGGAT CCATGATGTA 863
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   TGAGTGCTCT GACAAGGCTC CTACGGAGTC CGCCCGGCTC GTGCGGCACG AGTGAGGCTG CTGGCGGCGA 9863
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   CAAGTGCTCT GACAAGGCTC CTACGGAGTC CGCCCGGCTC GTGCGGCACG AGTGAGGCTG CTGGCGGCGA 934
CP7              TGAGTGCTCT GACAAGGCTC CTACGGAGTC CGCCCGGCTC GTGCGGCACG AGTGAGGCTG CTGGCGGCGA 1469
CP5              CAAGTGCTCT GACAAGGCTC CTACGGAGTC CGCCCGGCTC GTGCGGCACG AGTGAGGCTG CTGGCGGCGA 933
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
```

```
contig 0001218   CTGCTTGCCG GCTTCACGTA CCTCTCCCGA AACTCCAATT CTGTTTTGCT TTCATGCACT GTTCCAGTGT  9933
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521   CTGCTTGCCG GCTTCACGTA CCTCTCCCAA AACTCCAATT CTGTTTTGCT TTCATGCACT GTTCCAGTGT  1004
CP7              CTGCTTGCCG GCTTCACGTA CCTCTCCCGA AACTCCAATT CTGTTTTGCT TTCATGCACT GTTCCAGTGT  1539
CP5              CTGCTTGCCG GCTTCACGTA CCTCTCCCAA AACTCCAATT CTGTTTTGCT TTCATGCACT GTTCCAGTGT  1003
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1

contig 0001218   ATTTGCTTTC TTGGTGTGCG GCCACACGAT GCGGCTCTGA GGCGAGTGCG AGGGACACCG CGCCGCTCTG 10003
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521   ATTTGCTTTC TTGGTGTGCG GCCACACGAT GCGGCTCTGA GGCGAGTGCG AGGGACACCG CGCCGCTCTG  1074
CP7              ATTTGCTTTC TTGGTGTGCG GCCACACGAT GCGGCTCTGA GGCGAGTGCG AGGGACACCG CGCCGCTCTG  1609
CP5              ATTTGCTTTC TTGGTGTGCG GCCACACGAT GCGGCTCTGA GGCGAGTGCG AGGGACACCG CGCCGCTCTG  1073
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1

contig 0001218   CCTCTCAGCA CTTGCGGAGT GTGAGAAAGC TGCCCCAACC ACGCGTGTGT CTTTCTATGC TTTTGCTTGT 10073
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521   CCTCTCAGCA CTTGCGGAGT GTGAGAAAGC TGCCCCAACC ACGCGTGTGT CTTTCTATGC TTTTGCTTGT  1144
CP7              CCTCTCAGCA CTTGCGGAGT GTGAGAAAGC TGCCCCAACC ACGCGTGTGT CTTTCTATGC TTTTGCTTGT  1679
CP5              CCTCTCAGCA CTTGCGGAGT GTGAGAAAGC TGCCCCAACC ACGCGTGTGT CTTTCTATGC TTTTGCTTGT  1143
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1

contig 0001218   CCTTTTTCTC CATTACTTTT ACTGGCTCAC TGACCACGGC AACAGTCGCT GCTGCCACTG ACCCGCGAGG 10143
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521   CCTTTTTCTC CATTACTTTT ACTGGCTCAC TGACCACGGC AACAGTCGCT GCTGCCACTG ACCCGCGAGG  1214
CP7              CCTTTTTCTC ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5              CCTTTTTCTC ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1

contig 0001218   AGCGTCGCAC CCCGTTGGCC TTGTCTTTGT TGTTTTTGCG CCCACACTTG CTCCCAATAC GCTTCGCTTG 10213
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521   AGCGTCGCAC CCCGTTGGCC TTGTCTTT-T TATTTTTGCG CCCACACTTG CTCCCAATAC GCTTCGCTTG  1283
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1

contig 0001218   CCTGATTGCG CTTCTTCCTT TGCCTCCCGT CTCTCCCTGT ATGTCACACC GTACGT-TGG CGCAGTGTAT 10282
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521   CCTGATTGCG CTTCTTCCTT TGCCTCCCGT CTCTCCCTGT ATGTCACACC GTACGT-TGG CGCAGTGTAT  1352
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2              ---------- ---------- ---------- ---------- ------CAGT TTCTGTACTA TATTGTGTAT  24
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ----GTGTAT  6
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ----GTGTAT  6
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ----GTGTAT  6
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ----------  1
```


contig 0001218 CCTGCACTCA CGAACGAGCT AAAAATAAGA ATCAAAAAAG GGAGGACG-A TAGAGCAGCG ACAACCCGAG 10351
5' overhang 999
contig 0002521 CCTGCACTCA CGAACGAGCT AAAAATAAGA ATCAAAAAAG GGAGGACG-A TAGAGCAGCG ACAACCCGAG 1421
CP7 1689
CP5 1153
CP2 CCTGCACTCA CGAACGAGCT AAAAATAAGA ATCAAAAAAG GGAGAACGAA TAGAGCAGCG ACAACCCGAG 94
CP4 CCTGCACTCA CGAACGAGCT AAAAATAAGA ATCAAAAAAG GGAGAAC-AA TAGAGCAGCG ACAACCCGAG 75
CP3 CCTGCACTCA CGAACGAGCT AAAAATAAGA ATCAAAAAAG GGAGGACG-A TAGAGCAGCG ACAACCCGAG 75
CP8 CCTGCACTCA CGAACGAGCT AAAAATAAGA ATCAAAAAAG GGAGGACG-A TAGAGCAGCG ACAACCCGAG 75
CP1 1
CP6 1
contig 0001218 AAGGTAAGAG GCTTTTCTTC TTTCTTTTCC TTTATTGTAC GATGCCACGA TCAGAAATGA CACGCACCCT 10421
5' overhang 999
contig 0002521 AAGGTAAGAG GCTTTTCTTC TTTCTTTTCC TTTATTGTAC GATGCCACGA TCAGAAATGA CACGCACCCT 1491
CP7 1689
CP5 1153
CP2 AAGGTAAGAG GCTTTTCTTC TTTCTTTTCC TTTATTGTAC GATGCCACGA TCAGAAATGA CACGCACCCT 164
CP4 AGCGTAAGAG GCTTTTCTTC TTTCTTTTCC TTTATTGTAC GATGCCACGA TCAGAAATGA CACGCACCCT 145
CP3 AAGGTAAGAG GCTTTTCTTC TTTCTTTTCC TTTATTGTAC GATGCCACGA TCAGAAATGA CACGCACCCT 145
CP8 AAGGTAAGAG GCTTTTCTTC TTTCTTTTCC TTTATTGTAC GATGCCACGA TCAGAAATGA CACGCACCCT 145
CP1 1
CP6 1
contig 0001218 GCGCTTCTCC GTGGGCCTGC TCGCTGTTGC GGCGTGCTTT GTTCCCGTGG CGTTGGGAGT GCTTCACGCA 10491
5' overhang 999
contig 0002521 GCGCTTCTCC GTGGGCCTGC TCGCTGTTGC GGCGTGCTTT GTTCCCGTGG CGTTGGGAGT GCTTCACGCA 1561
CP7 1689
CP5 1153
CP2 GCGCTTCTCC GTGGGCCTGC TCGCCGTTGC GGCGTGCTTT GTTCCCGTGG CGTTGGGAGT GCTTCACGCA 234
CP4 GCGCTTCTCC GTGGGCCTGC TCGCCGTTGC GGCGTGCTTT GTTCCCGTGG CGTTGGGAGT GCTTCACGCA 215
CP3 GCGCTTCTCC GTGGGCCTGC TCGCTGTTGC GGCGTGCTTT GTTCCCGTGG CGTTGGGAGT GCTTCACGCA 215
CP8 GCGCTTCTCC GTGGGCCTGC TCGCTGTTGC GGCGTGCTTT GTTCCCGTGG CGTTGGGAGT GCTTCACGCA 215
CP1 1
CP6 1
contig 0001218 GAGCAATCGT TGCAGCAGCA ATTCGCCGCA TTCAAGCAAA AGTTCAGCAG GTCGTACAAG GACGCCACGG 10561
5' overhang 999
contig 0002521 GAGCAATCGT TGCAGCAGCA ATTCGCCGCA TTCAAGCAAA AGTACAGCAG GTCGTACAAG GACGCCACGG 1631
CP7 1689
CP5 1153
CP2 GAGCAATCGT TGCAGCAGCA ATTCGCCGCA TTCAAGCAAA AGTACAGCAG GTCGTACAAG GACGCCACGG 304
CP4 GAGCAATCGT TGCAGCAGCA ATTCGCCGCA TTCAAGCAAA AGTACAGCAG GTCGTACAAG GACGCCACGG 285
CP3 GAGCAATCGT TGCAGCAGCA ATTCGCCGCA TTCAAGCAAA AGTACAGCAG GTCGTACAAG GACGCCACGG 285
CP8 GAGCAATCGT TGCAGCAGCA ATTCGCCGCA TTCAAGCAAA AGTTCAGCAG GTCGTACAAG GACGCCACGG 285
CP1 1
CP6 1
contig 0001218 AGGAGGCATT CCGTTTCCGT GTCTTCAAGC AAAACATGGA GCGTGCAAAG GAGGAGGCCG CTGCGAACCC 10631
5' overhang 999
contig 0002521 AGGAGGCATT CCGTTTCCGC GTCTTCAAGC AGAACATGGA GCGTGCAAAG GAGGAGGCCG CTGCGAACCC 1701
CP7 1689
CP5 1153
CP2 AGGAGGCATT CCGTTTCCGT GTCTTCAAGC AGAACATGGA GCGTGCAAAG GAGGAGGCCG CTGCGAACCC 374
CP4 AGGAGGCATT CCGTTTCCGC GTCTTCAAGC AGAACATGGA GCGTGCAAA- GAGGAGGCCG CTGCGAACCC 354
CP3 AGGAGGCATT CCGTTTCCGC GTCTTCAAGC AGAACATGGA GCGTGCAAAG GAGGAGGCCG CTGCGAACCC 355
CP8 AGGAGGCATT CCGTTTCCGT GTCTTCAAGC AAAACATGGA GCGTGCAAAG GAGGAGGCCG CTGCGAACCC 355
CP1 1
CP6 1
contig 0001218 CTATGCGACG TTTGGTGTGA CGCGGTTCTC CGATATGTCA CCCGAGGAGT TCAGGGCGAC CTACCACAAC 10701
5' overhang 999
contig 0002521 CTATGCGACG TTTGGTGTGA CGCGGTTCTC CGATATGTCA CCCGAGGAGT TCAGGGCGAC CTACCACAAC 1771
CP7 1689
CP5 1153
CP2 CTATGCGACG TTTGGTGTGA CGCGGTTCTC CGATATGTCA CCCGAGGAGT TCAGGGCGAC CTACCACAAC 444
CP4 CTATGCGACG TTTGGTGTGA CGCGGTTTTC CGATATGTCA CCCGAGGAGT TCAGGGCGAC CTACCACAAC 424
CP3 CTATGCGACG TTTGGTGTGA CGCGGTTCTC CGATATGTCA CCCGAGGAGT TCAGGGCGAC CTACCACAAC 425
CP8 CTATGCGACG TTTGGTGTGA CGCGGTTCTC CGATATGTCA CCCGAGGAGT TCAGGGCGAC CTACCACAAC 425
CP1 1
CP6 1


contig 0001218 GGGGCGGAGT ACTACGCTGC GGCGCTGAAG CGACCACGCA AGGTGGTGAA TGTGTCCACT GGGAGGCCAC 10771
5' overhang 999
contig 0002521 GGGGCGGAGT ACTACGCTGC GGCGCTGAAG CGACCACGCA AGGTGGTGAA TGTGTCCACT GGGAAGGCAC 1841
CP7 1689
CP5 1153
CP2 GGGGCGGAGT ACTACGCTGC GGCGCTGAAG CGACCACGCA AGGTGGTGAA TGTGTCCACT GGGAAGGCAC 514
CP4 GGGGCGGAGT ACTACGCTGC GGCGCTGAAG CGACCACGCA AGGTGGTGAC TGTGTCCACT GGGAAGGCAC 494
CP3 GGGGCGGAGT ACTACGCTGC GGCGCTGAAG CGACCACGCA AGGTGGTGAA TGTGTCCACT GGGAAGGCAC 495
CP8 GGGGCGGAGT ACTACGCTGC GGCGCTGAAG CGACCACGCA AGGTGGTGAA TGTGTCCACT GGGAGGCCAC 495
CP1 1
CP6 1
contig 0001218 CGAT-GACGG TTGACTGGCG CAAGAAGGGC GCGGTGACAC CCGTGAAGGA CCAGGGGAAA TGCGACTCTT 10840
5' overhang 999
contig 0002521 CTGA-GGCAG TTGACTGGCG CAAGAAGGGC GCGGTGACAC CCGTGAAGGA CCAGGGGCAA TGCGGTTCTT 1910
CP7 1689
CP5 1153
CP2 CTGA-GGCAG TTGACTGGCG CAAGAAGGGC GCGGTGACAC CCGTGAAGGA CCAGGGGCAA TGCGGCTCTT 583
CP4 CTGA-GGCAG TTGACTGGCG CAAGAAGGGC GCGGTGACAC CCGTGAAGGA CCAGGGACAA TGCGGTTCTT 563
CP3 CTGA-GGCAG TTGACTGGCG CAAGAAGGGC GCGGTGACAC CCGTGAAGGA CCAGGGGCAA TGCGGTTCTT 564
CP8 C-GATGACGG TTGACTGGCG CAAGAAGGGC GCGGTGACAC CCGTGAAGGA CCAGGGGAAA TGCGACTCTT 564
CP1 1
CP6 1
contig 0001218 CGTGGGCATT CTCTGCCATA GGGAACATAG AGGGCCAGTG GAAGATTGCG GGCCATGAGC TGACGTCTTT 10910
5' overhang 999
contig 0002521 GCTGGGCATT CTCTGCCATA GGGAACATAG AGGGCCAGTG GAAGATTGCG GGCCATGAGC TGACGTCTCT 1980
CP7 1689
CP5 1153
CP2 GCTGGGCATT CTCTGCCATA GGGAACATAG AGGGCCAGTG GAAGGTTGCA GGCCATGAGC TGACGTCTTT 653
CP4 GCTGGGCATT CTCTGCCATA GGGAACATAG AGGGCCAGTG GAAGGTTACG GGCCATAATC TGACGTCTCT 633
CP3 GCTGGGCATT CTCTGCCATA GGGAACATAG AGGGCCAGTG GAAGATTGCG GGCCATGAGC TGACGTCTCT 634
CP8 CGTGGGCATT CTCTGCCATA GGGAACATAG AGGGCCAGTG GAAGATTGCG GGCCATGAGC TGACGTCTTT 634
CP1 1
CP6 1
contig 0001218 GTCGGAGCAA ATGCTCGTGT CATGTGACAC TAATGACCTT GGCTGCGAAC TTGGTCTCAA GGACCCCGCG 10980
5' overhang 999
contig 0002521 GTCGGAGCAG ATGCTCGTGT CATGTGACAC TAATGACTTT GGCTGCGGAG GTGGTCTCAT GGACGACGCG 2050
CP7 1689
CP5 1153
CP2 GTCGGAGCAG ATGCTCGTGT CATGTGACAC TAATGACTTT GGCTGCGAAG GTGGTCTCAT GGACGACGCG 723
CP4 GTCGGAGCAG ATGCTCGTGT CATGTGACAC TGAGGACCTT GGTTGCGCAG GTGGTCTCAT GGATAACGCG 703
CP3 GTCGGAGCAG ATGCTCGTGT CATGTGACAC TAATGACTTT GGCTGCGGAG GTGGTCTCAT GGACGACGCG 704
CP8 GTCGGAGCAA ATGCTCGTGT CATGTGACAC TAATGACCTT GGCTGCGAAC TTGGTCTCAA GGACCCCGCG 704
CP1 1
CP6 1
contig 0001218 TTCCAATGGA TTTTATGGTC GAATAAGGGC AATGTGTTCA CGGAGCAAAG TTACCCGTAC GCATCAGGAG 11050
5' overhang 999
contig 0002521 TTCAAATGGA TTGTGTCGTC GAATAAGGGC AATGTGTTCA CGGAGCAAAG TTACCCGTAC GCATCAGGAG 2120
CP7 1689
CP5 1153
CP2 TTCAAATGGA TTGTGTCGTC GAATAAGGGC AATGTGTTCA CGGAGCAAAG TTACCCGTAC GCATCAGGAG 793
CP4 TTCAAATGGA TTGTGTCGTC GAATAAGCAC AATGTGTTCA CGGAGGAAAG TTACCCGTAC GCCTCAAAAG 773
CP3 TTCAAATGGA TTGTGTCGTC GAATAAGGGC AATGTGTTCA CGGAGCAAAG TTACCCGTAC GCATCAGGAG 774
CP8 TTCCAATGGA TTTTATGGTC GAATAAGGGC AATGTGTTCA CGGAGCAAAG TTACCCGTAC GCATCAGGAG 774
CP1 1
CP6 1
contig 0001218 GGGGGAACGT GCCGACGTGC GACATGAGCG GCAAGGTCGT TGGTGCGAAG ATCAGTAACA TGCGCT-ACC 11119
5' overhang 999
contig 0002521 GGGGGAACGT GCCGGCGTGC GACATGAGCG GCAAGGTCGT TGGTGCGAAG ATCAGAGAC- CACGTTGACC 2189
CP7 1689
CP5 1153
CP2 GGGGGAACGT GCCGACGTGC GACAAGAGCG GCAAGGTCGT TGGTGCGAAG ATCAGAGAC- CACGTTGACC 862
CP4 GGGGGAACGT GCCGCCTTGC CGCATGAGCG GCAAAGTCGT TGGTGCGAAG ATCAGAGAC- CATGTTGACC 842
CP3 GGGGGAACGT GCCGGCGTGC GACATGAGCG GCAAGGTCGT TGGTGCGAAG ATCAGAGAC- CACGTTGACC 843
CP8 GGGGGAACGT GCCGACGTGC GACATGAGCG GCAAGGTCGT TGGTGCGAAG ATCAGTAACA TGCGCT-ACC 843
CP1 1
CP6 1

```
contig 0001218  TACCACTAGA CGAAGACACA ATAGCCGAGT GGCTGGCAAG GAAAGGGCCC GTCGCGATTG CCGTTGACGC 11189
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  TTCCGGAGGA CGAAAACGCA ATAGCCGAGT GGCTGGCAAA GAAAGGGCCC GTCGCGATTG CCGTTGACGC 2259
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             TTCCGGAGGA CGAAAACGCA ATAGCCGAGT GGCTGGCAAA GAACGGGCCC GTCGCGATTG CCGTTGACGC 932
CP4             TTCCGAAGGA CGAAAACGCA ATAGCCGAGT GGCTGGCAAA GAACGGGCCC GTCGCGATTG CCGTTGACGC 912
CP3             TTCCGGAGGA CGAAAACGCA ATAGCCGAGT GGCTGGCAAA GAAAGGGCCC GTCGCGATTG CCGTTGACGC 913
CP8             TACCACTAGA CGAAGACACA ATAGCCGAGT GGCTGGCAAG GAAAGGGCCC GTCGCGATTG CCGTTGACGC 913
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  CACCAGTTTC CAGAGGTACA CGGGCGGAGT GTTGACATCA TGCATCTCTC GA-CGACTCA ACTACGGTGC 11258
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  CACCAGTTTC CAGAGCTACA CGGGCGGAGT GTTGACATCA TGCATCTC-C GAGCATCTGG ACCATGGTGT 2328
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             CACCAGTTTC CAGAGCTACA CGGGCGGAGT GTTGACATCA TGCATCTC-C GAGCATCTGG ACCATGGTGT 1001
CP4             CACCAGTTTC CAGGACTACA CGGGCGGAGT GCTGACATCA TGCATCTC-T AAGCAACTGG ACCATGGTGT 981
CP3             CACCAGTTTC CAGAGCTACA CGGGCGGAGT GTTGACATCA TGCATCTC-C GAGCATCTGG ACCATGGTGT 982
CP8             CACCAGTTTC CAGAGGTACA CGGGCGGAGT GTTGACATCA TGCATCTCTC GA-CGACTCA ACTACGGTGC 982
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  GCTCCTCGTG GGGTATGACG ACACAAGCAA GCCACCATAC TGGATTATCA AGAACTCATG GGGCAAGGGA 11328
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  GCTCCTCGTG GGGTATGACG ACACAAGCAA GCCACCATAC TGGATTATCA AGAACTCATG GAGCAAGGGA 2398
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             GCTCCTCGTG GGGTATGACG ACACAAGCAA GCCACCATAC TGGATTATCA AGAACTCATG GAGCAAGGGA 1071
CP4             GCTCCTCGTT GGGTATGACG ACACGAGCAA GCCACCATAC TGGATTATCA AGAACTCATG GAGCGAGAAA 1051
CP3             GCTCCTCGTG GGGTATGACG ACACAAGCAA GCCACCATAC TGGATTATCA AGAACTCATG GAGCAAGGGA 1052
CP8             GCTCCTCGTG GGGTATGACG ACACAAGCAA GCCACCATAC TGGATTATCA AGAACTCATG GGGCAAGGGA 1052
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  TGGGGCGAGG AGGGTTACAT CCGCATTGAG AAGGGCACAA ACCAATGTCT TGTGAAAAAT CTCCCAAGCT 11398
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  TGGGGCGAGG AGGGTTACAT CCGCATTGAG AAGGGCACAA ACCAATGTCT TATGAAAAAT CTCCCAAGCT 2468
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             TGGGGCGAGG AGGGTTAC-T CCGCATTGAG AA-GGCAC-A ACCAATGTCT TATGAAAAAT CTCCCAAGCT 1138
CP4             TGGGGCGAGG AGGGTTACAT CCGCATTGAG AAGGGCACAA ACCAATGTCT TATGAAAAAT TACGCGACCT 1121
CP3             TGGGGCGAGG AGGGTTACAT CCGCATTGAG AAGGGCACAA ACCAATGTCT TATGAAAAAT CTCCCAAGCT 1122
CP8             TGGGGCGAGG AGGGTTACAT CCGCATTGAG AAGGGCACAA ACCAATGTCT TGTGAAAAAT CTCCCAAGCT 1122
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  CCGCCGTTGT CAGCGTCCCT CCGCCTCCAC CCCCGGCTCC GACCTTCACG CAGGAGTTGT GCGAGGGTGC 11468
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  CCGCCGTTGT CAGCGGCCCT CCGCCTCCGC CCCCGACTCC GACCTTCACG CAGGAGTTGT GCGAGGGTGC 2538
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             CCGCCGTTGT CAGCGGCCCT CCGCCTCCAC CCCCGACTCC GACCTTCACG CAGGAGTTGT GCGAGGGTGC 1208
CP4             CCGCCGTTGT CCACCGCCCC GTCCCTCCGC CTCCACCCCC GGCGTCGACG ---------- ---------- 1171
CP3             CCGCCGTTGT CAGCGGCCCT CCGCCTCCGC CCCCGACTCC GACCTTCACG CAGGAGTTGT GCGAGGGTGC 1192
CP8             CCGCCGTTGT CAGCGTCCCT CCGCCTCCAC CCCCGGCTCC GACCTTCACG CAGGAGTTGT GCGAGGGTGC 1192
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  CGAATGCTAG AGTAAGTGCA CCAAAGCCAC ATTCCCAACG GGCAAGTGCG TGCAGCTCAG CGGCGCCGGT 11538
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  CGAATGCCAG AGTAAGTGCA CCAAAGCCAC ATTCCCAACG GGCAAGTGCG TGCAGCTCAG CGGCGCCGGT 2608
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             CGAATGCCAG AGTAAGTGCA CCAAAGCCAC ATTCCCAACG GGCAAGTGCG TGCAGCTCAG CGGCGCCGGT 1278
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             CGAATGCCAG AGTAAGTGCA CCAAAGCCAC ATTCCCAACG GGCAAGTGCG TGCAGCTCAG CGGCGCCGGT 1262
CP8             CGAATGCTAG AGTAAGTGCA CCAAAGCCAC ATTCCCAACG GGCAAGTGCG TGCAGCTCAG CGGCGCCGGT 1262
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
```

```
contig 0001218  TCTGTCATCG CCTCGTGCGG CTCCAATAAC CTCACGCAGA TCGTCTACCC GTTGAGCAGC TCCTGCAGCG 11608
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  TCTGTCATCG CCTCGTGCGG CTCCAATAAC CTCACGCAGA TCGTCTACCC GTTGAGCAGC TCCTGCAGCG 2678
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             TCTGTCATCG CCTCGTGCGG CTCCAATAAC CTCACGCAGA TCGTCTACCC GTTGAGCAGC TCCTGCAGCG 1348
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             TCTGTCATCG CCTCGTGCGG CTCCAATAAC CTCACGCAGA TCGTCTACCC GTTGAGCAGC TCCTGCAGCG 1332
CP8             TCTGTCATCG CCTCGTGCGG CTCCAATAAC CTCACGCAGA TCGTCTACCC GTTGAGCAGC TCCTGCAGCG 1332
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  GCTTCTCCGT TCCGTTGACT GTGCCACTGG ACAAGTGCCT GCCCATTGTG ATTGGATCCG TGATGTATGA 11678
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  GCTACTCCGT TCCGTTGACT GTGCCACTGG ACAAGTGCCT GCCCATCGTG ATTGGATCCG TGATGTATGA 2748
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             GCTTCTCCGT TCCGTTGACT GTGCCACTGG ACAAGTGCCT GCCCATTGTG ATTGGATCCG TGATGTATGA 1418
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             GCTACTCCGT TCCGTTGACT GTGCCACTGG ACAAGTGCCT GCCCATCGTG ATTGGATCCG TGATGTATGA 1402
CP8             GCTTCTCCGT TCCGTTGACT GTGCCACTGG ACAAGTGCCT GCCCATTGTG ATTGGATCCG TGATGTATGA 1402
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1


contig 0001218  GTGCTCTGAC AAGGCTCCTA CGGAATCCGC CCGGCTCGTG CGGCACGAGT GAGGCTGCTG GCGGCGACTG 11748
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  GTGCTCTGGC AAGGCTCCTA CGGAATCCGC CCGGCTCGTG CGGCACGAGT GAGGCTGCTG GCGGCGACTG 2818
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             GTGCTCTGAC AAGGCTCCTA CGGAATCCGC CCGGCTCGTG CGGCACGAGT GAGGCTGCTG GCGGCGACTG 1488
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             GTGCTCTGGC AAGGCTCCTA CGGAATCCGC CCGGCTCGTG CGGCACGAGT GAGGCTGCTG GCGGCGACTG 1472
CP8             GTGCTCTGAC AAGGCTCCTA CGGAATCCGC CCGGCTCGTG CGGCACGAGT GAGGCTGCTG GCGGCGACTG 1472
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  CTTGCCGGCT TCACGTACCT CTCCCGAAAC TCCAATTCTG TTTTGCTTTC ATGCACTGTT CCAGTGTATT 11818
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  CTTGCCGGCT TCACGTACCT CCCCCAAAAC TCCAATTCTG TTTTGCTTTC ATGCACTGTT CCAGTGTATT 2888
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             CTTGCCGGCT TCACGTACCT CTCCCGAAAC TCCAATTCTG TTTTGCTTTC ATGCACTGTT CCAGTGTATT 1558
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             CTTGCCGGCT TCACGTACCT CCCCCAAAAC TCCAATTCTG TTTTGCTTTC ATGCACTGTT CCAGTGTATT 1542
CP8             CTTGCCGGCT TCACGTACCT CTCCCGAAAC TCCAATTCTG TTTTGCTTTC ATGCACTGTT CCAGTGTATT 1542
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  TGCTTTCTTG GTGTGCGGCC ACACAATGCG GCTCTGAGGC GAGTGCGAGG GACACCGCGC CGCTCTGCCT 11888
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  TGCTTTCTTG GTGTGCGGCC ACACGATGCG GCTCTGAGGC GAGTGCGAGG GACACCGCGC CGCTCTGCCT 2958
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             TGCTTTCTTG GTGTGCGGCC ACACAATGCG GCTCTGAGGC GAGTGCGA-G GACA-CGCGC CGCTCTGCCT 1626
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             TGCTTTCTTG GTGTGCGGCC ACACGATGCG GCTCTGAGGC GAGTGCGAGG GACACCGCGC CGCTCTGCCT 1612
CP8             TGCTTTCTTG GTGTGCGGCC ACACAATGCG GCTCTGAGGC GAGTGCGAGG GACACCGCGC CGCTCTGCCT 1612
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218  CTCAGCACTT GCGGAGTGTG AGAAAGCTGC CCCAACCACG CGTGTGTCTT TCTATGCTTT TGCTTGTCCT 11958
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  CTCAGCACTT GCGGAGTGTG AGAAAGCTGC CCCAACCACG CGTGTGTCTT TCTATGCTTT TGCTTGTCCT 3028
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             CTCAGCACTT GCGGAGTGTG AGAAAGCTGC CCCAACCACG CGTGTGTCTT TCTATGCTTT TGCTTGTCCT 1696
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             CTCAGCACTT GCGGAGTGTG AGAAAGCTGC CCCAACCACG CGTGTGTCTT TCTATGCTTT TGCTTGTCCT 1682
CP8             CTCAGCACTT GCGGAGTGTG AGAAAGCTGC CCCAACCACG CGTGTGTCTT TCTATGCTTT TGCTTGTCCT 1682
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
```

```
contig 0001218   TTTTCTCCAT TACTTTTACT GGCTCACTGG CCACGGCAAC AGTCGCTGCT GCCACTGACC CGCGAGGAGC 12028
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   TTTTCTCCAT TACTTTTACT GGCTCACTGA CCACGGCAAC AGTCGCTGCT GCCACTGACC CGCGAGGAGT 3098
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              TTTTCTCG-- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              TTTTCTC--- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              TTTTCTC--- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   GTCGCACCCC GTTGGCCTTG TCTTT-TTAT TTTTGCACCC ACACTTGCTC CCAATACGCT TCGCTTGCCT 12097
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   GTCGCACCCC GTTGGCCTTG TTTTTGTTAT TTTTGCACCC ACACTTGCTC CCAATACGCT TCGTTTGCCT 3168
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   GATTACGCTC CCTCCTTTGC CTCCCGTCTC TCCCTGTATG TCACACCGTA CGTTGGCGCA GTGTATCCTG 12167
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   GATTGCGCTT CTTCCTTTGC CTCCCGTCTC TCCCTGTATG CCACACCGTA CGTTGGCGCA GTGTATCCTG 3238
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   CACTCACGAA CGAGCTAAAA ATAAGAATCA AAAAAGGGAG AACAATAGAG CAGCGACAAC CCGAGAAGGT 12237
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   CACTCACGAA CGAGCTAAAA ATAAGAATCA AAAAAGGGAG AACAATAGAG CAGCGACAAC CCGAGAGCGT 3308
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              ---------- ---------- ---------A AAAAAGGGAG AACAATAAAG TAGCGACAAC CAGAGAAGGT 41
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   AAGAGGCTTT TCTTCTTTCT TTTCCTTTAT TGTACGATGC CACGATCAGA AATGACACGC ACCCTGCGCT 12307
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   AAGAGGCTTT TCTTCTTTCT TTTCCTTTAT TGTACGATGC CACGATCAGA AATGACACGC ACCCTGCGCT 3378
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              GAGAGGCCTT TCTTCTTTCT TTTCCTTTAT TGTACGATGC CACGATCAGA AATGACACGC ACCCTGCGCT 111
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   TCTCCGTGGG CCTGCTCGCT GTTGCGGCGT GCTTTGTTCC CGTGGCGTTG GGAGTGCTTC ACGCAGAGCA 12377
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   TCTCCGTGGG CCTGCTCGCC GTTGCGGCGT GCTTTGTTCC CGTGGCGTTG GGAGTGCTTC ACGCAGAGCA 3448
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              TCTCCGTGGG CCTGCTCGCC GTCGCGGCGT GCTTTGTTCC CGTGGCGTTG GGAGTGCTTC ACGCAGAGCA 181
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
```

```
contig 0001218   ATCGTTGCAG CAGCAATTCG CCGCATTCAA GCAAAAGTAC AGCAGGTCGT ACAAGGACGC CACGGAGGAG 12447
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   ATCGTTGCAG CAGCAATTCG CCGCATTCAA GCAAAAGTAC AGCAGGTCGT ACAAGGACGC CACGGAGGAG 3518
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              ATCGTTGCAG CAGCAATTCG CCGCATTCAA GCAAAAGTAC AGCAGGTCGT ACAAGGACGC CACGGAGGAG 251
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   GCATTCCGTT TCCGTGTCTT CAAGCAGAAC ATGGAGCGTG CAAAGGAGGA GGCCGCTGCG AACCCCTATG 12517
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   GCATTCCGTT TCCGCGTCTT CAAGCAGAAC ATGGAGCGTG CAAA-GAGGA GGCCGCTGCG AACCCCTATG 3587
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              GCATTCCGTT TCCGCGTCTT CAAGCAGAAC ATGGAGCGTG CAAAGGAGGA GGCCGCTGCG AACCCCTATG 321
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   CG-------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   CGACGTTTGG TGTGACGCGG TTTTCCGATA TGTCACCCGA GGAGTTCAGG GCGACCTACC ACAACGGGGC 3657
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              CGACGTTTGG TGTGACGCGG TTCTCCGATA TGTCACCCGA GGAGTTCAGG GCGACCTACC ACAACGGGGC 391
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   GGAGTACTAC GCTGCGGCGC TGAAGCGACC ACGCAAGGTG GTGACTGTGT CCACTGGGAA GGCACCTGAG 3727
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              GGAGTACTAC GCTGCGGCGC TGAAGCGGCC ACGCAAGGTG GTGAATGTGT CCACTGGGAA GGCACCACCG 461
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   GCAGTTGACT GGCGCAAGAA GGGCGCGGTG ACACCCGTGA AGGACCAGGG ACAATGCGGT TCTTGCTGGG 3797
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              GCAGTTGACT GGCGCAAGAA GGGCGCAGTG ACACCCGTGA AGGACCAGGG GGCATGCGGT TCTTGCTGGG 531
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1

contig 0001218   ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521   CATTCTCTGC CATAGGGAAC ATAGAGGGCC AGTGGAAGGT TACGGGCCAT AATCTGACGT CTCTGTCGGA 3867
CP7              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1              CGTTCTCTGC CATAGGGAAC ATAGAGGGCC AGTGGAAGGT TGCGGGCCAT GAGCTGACGT CTCTGTCGGA 601
CP6              ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1
```

```
contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  GCAGATGCTC GTGTCATGTG ACACTGAGGA CCTTGGTTGC GCAGGTGGTC TCATGGATAA CGCGTTCAAA  3937
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             GCAAATGCTC GTGTCATGTG ACACTACGGA CTATGGCTGC CGAGGAGGCC TCATGGACAA ATCGCTTCAA  671
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  TGGATTGTGT CGTCGAATAA GCACAATGTG TTCACGGAGG AAAGTTACCC GTACGCCTCA AAAGGGGGGA  4007
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             TGGATTGTGT CGTCGAATAA GGGCAATGTG TTCACGGCGC AAAGTTACCC GTACGCATCA GGAGGGGGGA  741
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  ACGTGCCGCC TTGCCGCATG AGCGGCAAAG TCGTTGGTGC GAAGATCAGA GACCATGTTG ACCTTCCGAA  4077
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             AGATGCCGCC TTGTAACAAG AGCGGCAAAG TCGTTGGTGC AAAGATCAGT GGCCACATTA ACCTTCCGAA  811
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  GGACGAAAAC GCAATAGCCG AGTGGCTGGC AAAGAACGGG CCCGTCGCGA TTGCCGTTGA CGCCACCAGT  4147
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             GGACGAAAAC GCAATAGCGG AGTGGCTGGC AAAGAACGGG CCCGTCGCGA TTGCCGTTGA CGCCACCAGT  881
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  TTCCAGGACT ACACGGGCGG AGTGCTGACA TCATGCATCT CTAAGCAACT GGACCATGGT GTGCTCCTCG  4217
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             TTCCTGGGCT ACAAAGGCGG AGTGTTGACA TCATGCATCT CTAAGGGACT GGATCATGAT GTGCTCCTCG  951
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  TTGGGTATGA CGACACGAGC AAGCCACCAT ACTGGATTAT CAAGAACTCA TGGAGCGAGA AATGGGGCGA  4287
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             TGGGGTATGA CGACACGAGC AAGCCACCAT ATTGGATTAT CAAGAACTCA TGGAGCAAGG GATGGGGCGA  1021
CP6             ---------- ---------- ---------- ---------- ---------- ------GAGA AATGGGGCGA  14
```

```
contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  GGAGGGTTAC ATCCGCATTG AGAAGGGCAC AAACCAATGT CTTATGAAAA ATTACGCGAC CTCCGCCGTT 4357
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             GGAGGGTTAC ATCCGCATTG AGAAGGGCAC AAACCAATGT CTCATGAAAA ATTACGCGAG GTCAGCCGTT 1091
CP6             GGAGGGTTAC ATCCGCATTG AGAAGGGCAC AAACCAATGT CTTATGAAAA ATTACGCGAC CTCCGCCGTT 84

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  GTCCACCGCC CCGTC-CCTC CGCCTCCACC CCCGGCGTCG ACGTTCACGC AGGAGTTCTG CGAGGGTGCC 4426
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             GT-CAGCGGC CCTCCGCCTC CGCCTCCACC CCCGGCGTCG ACCTTCACGC AGGAGTTCTG CGAGGGTGCA 1160
CP6             GTCCACCGCC CCGTC-CCTC CGCCTCCACC CCCGGCGTCG ACGTTCACGC AGGAGTTCTG CGAGGGTGCC 153

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  GAATGCCAGA GTGGCTGCAC CAAAGCCACA TTCCCAACGG GCAAGTGCGT GCAGCTCAGC GGCGCCGGTT 4496
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             GAGTGCCAGA GTGGTTGCAC CAAAGCCACA TTCCCAACGG GCAAGTGCGT GCAGTTCGGC GGCGCCGGTT 1230
CP6             GAATGCCAGA GTGGCTGCAC CAAAGCCACA TTCCCAACGG GCAAGTGCGT GCAGCTCAGC GGCGCCGGTT 223

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  CTGTCATCGC CTCGTGCGGC TCCAATAACC TCACACAGAT CGTCTACCCG TTGAGCAGCT CCTGCAGCGG 4566
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             CTGTCATCGC CTCGTGCGGC TCCAATAACC TCACACAGAT CGTCTACCCG TTGAGCAGCT CCTGCAGCGG 1300
CP6             CTGTCATCGC CTCGTGCGGC TCCAATAACC TCACACAGAT CGTCTACCCG TTGAGCAGCT CCTGCAGCGG 293

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  CTACTCCGTT CCGTTGACTG TGCCACTGGA CAAGTGCCTG CCCATCGTGA TTGGATCCGT GATGTATGAG 4636
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             CTTCTCCATT CCGTTGACTG TGCCACTGGA CAAGTGCCTG CCCATCGTGG TTGGATCCGT GATGTATGAG 1370
CP6             CTACTCCGTT CCGTTGACTG TGCCACTGGA CAAGTGCCTG CCCATCGTGA TTGGATCCGT GATGTATGAG 363

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  TGCTCTGGCA ATGCTCCTAC GGAATCCGCC CGGCTCGTGC GACACGAGTG AGGCTGCTGG CGGCGACTGC 4706
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             TGCTCTGGCA AGGCTCCTAC GGAATCCGCC CGGCTCGTGC GGCACGAGTG AGGCTGCTGG CGGCGACTGC 1440
CP6             TGCTCTGGCA ATGCTCCTAC GGAATCCGCC CGGCTCGTGC GACACGAGTG AGGCTGCTGG CGGCGACTGC 433
```

```
contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  TTGCCGGCTT CACGTACCTC TCCCAAAACT CCAATTCTGT TTTACTTTCA TGCATTGTTC CAGTGTATTT 4776
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             TTGCCGGCTT CACGTACCTC TCCCAAAACT CCAATTCTGT TTTGCTTTCA TGCACTGTTC CAGTGTATTT 1510
CP6             TTGCCGGCTT CACGTACCTC TCCCAAAACT CCAATTCTGT TTTACTTTCA TGCATTGTTC CAGTGTATTT 503

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  GCTTTCTTGG TGTGCGGCCA CACGATGCGG CTCTGAGGCG AGTGCGAGGG ACACCGCGCC GCTCTGCCTC 4846
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             GCTTTCTCGG TGTGCGGCCA CACAATGCGG CTCTGAGGCG AGTGCGAGGG ACA-CGCGCC GCTCTGCCTC 1579
CP6             GCTTTCTTGG TGTGCGGCCA CACGATGCGG CTCTGAGGCG AGTGCGAGGG ACACCGCGCC GCTCTGCCTC 573

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  TCAGCACTTG CGGAGTGTGA GAAAGCTGCC CCAACCACGC GTGTGTCTTT CTATGCTTTT GCTTGTCCTT 4916
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             TCAGCACTTG CGGAGTGTGA GAAAGCTGCC CCAACCACGC GTGTGTCTTT CTATGCTTTT GCTTGTCCTT 1649
CP6             TCAGCACTTG CGGAGTGTGA GAAAGCTGCC CCAACCACGC GTGTGTCTTT CTATGCTTTT GCTTGTCCTT 643

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  TTTGTCCATT ACTTTTACTG GCTCACTGGC CACGACAACA GTTGGTGCTG CCACTGACCC GCGAGTGGTG 4986
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             TTTCACCATT ACTTTTATTG GCTCACTGGC CACGGCCACA GTCGCTGCTG CCACTGACCC GCGAGGGGCG 1719
CP6             TTTCTC---- ---------- ---------- ---------- ---------- ---------- ---------- 649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  TCGCACCCCG TTGGCCTTGT TTTTTATTTT GTGAAAGCGT CTGGTATAAA TGTTTTAGGA CTTGCCCTCT 5056
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             TCACACCCCG TCGGCCTTGT TTTTTG---- ---------- ---------- ---------- ---------- 1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ---------- 12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 999
contig 0002521  CGTGACAAAT GTCTGTTATT CTCTTTCAGC TATTTGGAAA ACGTGTGCGT TTTTGAAATG CAGGTTTATC 5126
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ---------- 649
```

```
contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  TAAACAGAGT TTATGATATC TACACGGCCG TCCCTCCCTC AGGTGGAATG GGCAACGAGT TCCATGGCAT  5196
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  TTTTGGATTA CTTTCGCTGA TGTTAAGGAA CTTCCTTTTT GTCGTGCTTT CGACCGGCAG AGGACCTGAG  5266
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  AGTCACGTAT CATGTGGACG GGCTTCTAAC GATTGAGGAT GGAAATGTCA TGCTTTACTT TGAGCGCGGA  5336
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  CATAAAATGG ATTTCGTATC ATTAAGGCAG TGGATGGTAC GTTTTACGCT GCCCAGAATC TCAGATTTTC  5406
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  TTTGAGTGTT TGAAGAGGCG AGAGAATCAC TCCGTGCGTA ATGGACGACT GTAGAAGCCA CGGAATCTGC  5476
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  ACTCCATAAT CAATATGAGA GTGCTTGCGC GCAGCTCCCT GACGAGAGCT TCCACGTGCA GCAAAATGCG  5546
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649
```

```
contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  AATGAGCGGT TGGGCGGGAT GCAGCAACAG CTTGAACATC TGCAAGGAAT TATTGCACAA CAAGCAGCAC  5616
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  AGTTAGAAGA ATATGTCCGA ACGGATGCTA CGTTACGGCA ACAATTGACT GGTCAAGGTG TAGGCGTTCA  5686
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  CACAGAGGAG GGTGAAAGGC ACCATGGAAA GCAGGACACG GGTAATGTAA AACGTATGGA TGAATACGCT  5756
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  TGGAAAGGGT TCCCACCGCA CGCGCCTCGA CCTTCATCAG GCGGTGTTGG CATCCGGGGA ATAAAGGAGT  5826
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  TACAGGGGAA CTCCTGGATG TCAGGAGTGA GACCCGAATA TAAATATTCG TACATTAGTG AGTGCGATGG  5896
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  AGTATCTTGA GGACAGAAAT CCGGAACAAC TGATCGCGGC GTATCAGTAC ATGTCTCCGA CAGCGGCAAC  5966
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649
```

```
contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  ACCGGGACCG GCTTATACGG AGGCTTGGAA AAACTTCGTA TCGTGGGCAT TGGGGTTGCG CTCGGAGGGC  6036
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  ATATCGGAGG GGCACCTCCA ATTGGGGAGG CAGTTGTATA ACCAACTGCG GACAATACAG ACAGTTAAAG  6106
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  ATCACCCAGG TGTGAAGATT GAGGACATAA GAGAGTCAGT TGAGTACCCT CATAGAAAGG ATGATGTTTT  6176
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649
contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  CATTCAAGCG GCTCTGAAGG CAAAAAGGCA TCAAACCGCG ATCCCTCCAG GTCGGGCGCA GTGTTTCACA  6246
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  TGCGGTCGCG TGGTGCACTA CAGCAGCTCG TGTCGTGTAT CCAAGAAGAA ACCAGAGAGT CACCCAACAA  6316
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  ACGGTTTCAG GCCCCGGAGG CAATGACTTC GGGGGCCTTT AATACCTTTT CCCATACCAC ACAGGGGAAG  6386
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649
```

```
contig 0001218  ---------- ---------- ---------- ---------- ---------- ---------- ----------  12519
5' overhang     ---------- ---------- ---------- ---------- ---------- ---------- ----------  999
contig 0002521  CAGAAAAGAG GGTGTCATGT GGTGCAATAT TTCTCGTGTT CGGGCGAGTC CCCCGACACA GACGGGGAGA  6456
CP7             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP5             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1153
CP2             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1704
CP4             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1171
CP3             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP8             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1689
CP1             ---------- ---------- ---------- ---------- ---------- ---------- ----------  1745
CP6             ---------- ---------- ---------- ---------- ---------- ---------- ----------  649

contig 0001218  ---------- ---------- ---------- ---------- - 12519
5' overhang     ---------- ---------- ---------- ---------- - 999
contig 0002521  GTGTGGCCGT TACGGATAGC AAAGTCAGCA GCTGCTCAAT G 6497
CP7             ---------- ---------- ---------- ---------- - 1689
CP5             ---------- ---------- ---------- ---------- - 1153
CP2             ---------- ---------- ---------- ---------- - 1704
CP4             ---------- ---------- ---------- ---------- - 1171
CP3             ---------- ---------- ---------- ---------- - 1689
CP8             ---------- ---------- ---------- ---------- - 1689
CP1             ---------- ---------- ---------- ---------- - 1745
CP6             ---------- ---------- ---------- ---------- - 649
```

Appendix 3: Multiple sequence alignment of the Oligopeptidase B gene in different trypanosome species. Residues in colour were used to design primers in *Tb*OPB and *Tco*OPB for RNAi.

Key: (***** = identical), (**.** = very similar), (**:** = similar).

```
T. brucei          MQTERGPIAAHRPHEVVFGKVEGEDRGANPMDPPRRRVDPLFWLRDDNRADPEVLAHLHL 60
T. evansi          MQTERGPIAAHRPHEVVFGKVEGLDRGANPMDPPRRRVDPLFWLRDDNRADPEVLAHLHL 60
T. congolense      MTSDRGPIAAHKPYEVVFGKVEGEDRGPNPMDPPRRRVDPLFWLRDDSRTNPDVLAHLHL 60
T. vivax           MPAPCGPVAEQQDHEVIFGKVEGEDRGPNPMDPPLRRNDPLFWLRDDNRTNPKVLAHLHL 60
T. cruzi           --MKCGPIATPKDHEVVFGYVEGENRGNNAMNPPRRRNDPLFWLRDDSRKNPEVIAHLKL 58
                        **:*  : :**:** *** :** *.*:** ** *********.* :*.*:***:*

T. brucei          EKDYYEKRAVDIKDLAETIYQEHISHIEETDMSAPYVYDRFLYYTRDVKGLSYKLHCRVP 120
T. evansi          EKDYYEKRAVDIKDLAETIYQEHISHIEETDMSAPYVYDRFLYYTRDVKGLSYKLHCRVP 120
T. congolense      ERDYFEKRTGDIKDLAETIYQEHISHIEETDMSPPYTYNRFVYYTRRVKGLSYKIHCRVP 120
T. vivax           EKAYFEECTVDLKDLSETVYQEHLSHIQETDMSAPYVHDNYMYYKREVKGLSYKIHCRVP 120
T. cruzi           EQAYFEERTADIKDFSETIFKEYFSHIKETDISAPYRYDGYTYYTREVEGLSYKIHCRVP 118
                   *: *:*: : *:**::**:::*::***:***:*.** :: : **.* *:*****:*****

T. brucei          AGKTPGEGEDEEIVLDENKLAEGKSFCVVGCVAPAPPEHALVAYSVDYCGDEVYSIRFVR 180
T. evansi          AGKTPGEGEDEEIVLDENKLAEGKSFCVVGCVAPAPPEHALVAYSVDYCGDEVYSIRFVR 180
T. congolense      LGKVPGEGPDEEIVLDENKLAEGKAFCDVRSVAPAPPEHMLVAYSVDHLGDELYSIQFVG 180
T. vivax           LGKTPG-SSDEQVVLDENKLAEGKAFCHVRQVEPAPPKHSLVAYSVDYTGNEVYTIRFVG 179
T. cruzi           LGMTPGKSEDEQVILDENKLAEGKSFCMVHEVKPAPPDHNLVAYSVDYVGNEMYTIRFTD 178
                    * .**.. **:::**********:** *  * ****.* *******: *:*:*:*:*.

T. brucei          DVVADKVEGTNGSVVWGPNAECFFYITKDASKRDNKVWRHIIGQPQSEDVCLYTDDDPLF 240
T. evansi          DVVADKVEGTNGSVVWGPNAECFFYITKDASKRDNKVWRHIIGQPQSEDVCLYTDDDPLF 240
T. congolense      DASPDKLEGTTGSIIWGTNAECFFYVTPDSTKRSNKVWRHIIGQSQSEDVCLYTDDDPLF 240
T. vivax           DDVPDVVEGTNGHIVWGPDGGCFFYTTKDAAQRDYKVWRHVIGQQQSEDVCLYTEVDAVF 239
T. cruzi           NAVTDVVEGTNGQILWGPNASCFFYITKDAAERNYKIWRHIIGRPQSEDVCLYTENDLLF 238
                   :  .* :***.* ::**.:. **** * *:::*. *:***:**: *********: * :*

T. brucei          SVGVGRSGDGKTLIICSMSSETSESHLLDLRKGVKHNTLEMVRPREKGVRYTVEMHGTDT 300
T. evansi          SVGVGRSGDGKTLIICSMSSETSESHLLDLRKGVKHNTLEMVRPREKGVRYTVEMHGTDT 300
T. congolense      SVAASKSGDGHTLLISSSSSETTELHLLDLRKGLNNTTLEVVRKREKNVRYEVEMHGTET 300
T. vivax           SACMSKSGDGNTLLITSSSSETTEVHLLDLRGGVAHNELETVRPREKGVRYDVALHGTDT 299
T. cruzi           STAMAISGDGNTLIIGSGSFETTESHLLDLRKGNGHTTLEMVRPRERGVRYDVELHGTET 298
                   *.  . ****:**:* * * **:* ****** *  :. ** ** **:.*** * :***:*

T. brucei          LIVLTNKDKCVNGKVVLTKRSAPTDWGTVLIPHDDKVTIDDVAVFAKFAVLSGRRDGLTR 360
T. evansi          LIVLTNKDKCVNGKVVLTKRSAPTDWGTVLIPHDDKVTIDDVVVFAKFAVLSGRRDGLTR 360
T. congolense      LVILTNKDKCINGKVVLAKRASPSEWTNVLVPHDDKVFIDDIAVFAKFAVLSGRRDGLTR 360
T. vivax           LLVLTNKDKCINGQVLIVQRSAPSDWTRVLVPHSEEVFIEEIAVFSTFAVLAGRRAGLSR 359
T. cruzi           LLILTNKDNCMNGKVVSATRNTPSDWENVIVPHSEDVFIGEIGVFAKFVVLSGRRGGITR 358
                   *::*****:*:**:*: . * :*::*  *::**.:.* * :: **:.*.**:*** *::*

T. brucei          VWTVRLGPDNLFSS-ATLKELHFDEPVFTAHVVCSQMKTYDASLLRLRYSSMTTPTVWYD 419
T. evansi          VWTVRLGPDNLFSS-ATLKELHFDEPVFTAHVVCSQMKTYDASLLRLRYSSMTTPTVWYD 419
T. congolense      VWTVQVGPDNCFSA-GTLRELQFDEPVFTAHVITSQMKTYDTSSLRLEYSSMTTPTTWFD 419
T. vivax           IWTMQVGPDNTFSN-SLVNELQFDEPVFTTHVVLSQMHMYNTATLRLTYSSMTTPTTWYD 418
T. cruzi           VWAMPVGSDGLFRSGAFVQEVSFDEPVFTAFPVFSHMKMYDTETLRVSYTSMSTPTTWFD 418
                   :*:: :*.*. *   . :.*: *******:. : *:*: *::  **: *:**:***.*:*

T. brucei          EDVLSGERKVVKARKVGGGFESKNYVCRRELATAPDGTKVPISLVYDTSIDLKKPNPTML 479
T. evansi          EDVLSGERKVVKARKVGGGFESKNYVCRRELATAPDGTKVPISLVYDTSIDLKKPNPTML 479
T. congolense      EDLASGKRTVVKVSKVGGGFDSKNYVCQRRLATAPDGTTIPLSILYDVSLDMKKPHPTML 479
T. vivax           VHVAEGGRTPVKIREVGGGFNPAHYVSKRLLATAPDGTKIPISLVYDAALDLTKPHPTML 478
T. cruzi           LHVVNGTRTIVKVREVLGNFDAKNYTCRRLFATAPDRTKIPLSIVHDVSLDMSKPHPTVL 478
                    .: .* *. **  :* *.*:. :*..:* :***** *.:*:*:::*.::*:.**:**:*
```

```
T. brucei          YGYGSYGICIEPEFNSRFLPYVDRGMIYAIAHVRGGGEMGRTWYEVGGKYLTKRNTFMDF 539
T. evansi          YGYGSYGICIEPEFNSRFLPYVDRGMIYAIAHVRGGGEMGRTWYEVGGKYLTKRNTFMDF 539
T. congolense      YGYGSYGICVEPQFDIRCLPYVDRGVIYAIAHVRGGGEMGRAWYEIGGKYLTKRNTFMDF 539
T. vivax           YGYGSYGICVEPEFNIQYLPYVDRGVIFAIAHVRGGGEMGRAWYELGAKYLTKRNTFMDF 538
T. cruzi           YAYGSYGACVEPEFSVKYLPYLDRGVIYVIAHVRGGGEMGRAWYEVGAKYLTKRNTFSDF 538
                   *.***** *:**:*. : ***:***:*:.***********:***:*.********* **

T. brucei          IACAEHLISSGLTTPAQLSCEGRSAGGLLVGAVLNMRPDLFHVALAGVPFVDVMTTMCDP 599
T. evansi          IACAEHLISSGLTTPAQLSCEGRSAGGLLVGAVLNMRPDLFHVALAGVPFVDVMTTMCDP 599
T. congolense      ISCAEHLISSGVTTPPQLACEGRSAGGLLVGAVLNMRPDLFRVAVAGVPFVDVMTTMCDP 599
T. vivax           IACAEHLISSGLTTPNQLACEGRSAGGLLIGAVLNMRPDLFHVALAGVPFVDVMTTMCDP 598
T. cruzi           IACAEYLIEIGLTTPSQLACEGRSAGGLLIGAVLNMRPDLFRVALAGVPFVDVMTTMCDP 598
                   *:***:**. *:*** **:**********:***********:**:***************

T. brucei          SIPLTTGEWEEWGNPNEYKFFDYMNSYSPIDNVRAQDYPHLMIQAGLHDPRVAYWEPAKW 659
T. evansi          SIPLTTGEWEEWGNPNEYKFFDYMNSYSPIDNARAQDYPHLMIQAGLHDPRVAYWEPAKW 659
T. congolense      SIPLTTGEWEEWGNPNEYKFFDYMNSYSPIDNVRPQDYPNLIIQAGLHDPRVAYWEPAKW 659
T. vivax           TIPLTTFEWEEWGNPNEYKYFDYMNSYSPIDNVRAQAYPHLMIQAGLHDPRVAYWEPVKW 658
T. cruzi           SIPLTTGEWEEWGNPNEYKFFDYMNSYSPVDNVRAQDYPHLMIQAGLHDPRVAYWEPAKW 658
                   :***** ***********:*********:**.*.* **:*:***************.**

T. brucei          ASKLRELKTDSNEVLLKMDLESGHFSASDRYKYLRENAIQQAFVLKHLNVRQLLRK 715
T. evansi          ASKLRELKTDSNEVLLKMDLESGHFSASDRYKYLRENAIQQAFVLKHLNVRQLLRK 715
T. congolense      ASKLRELKTDNNEVLLKMDLDSGHFSASDRYKYLREHAIQQAFVLKHLGVRRLLRH 715
T. vivax           ASRLRQLKTDGNEVLVKMDLDSGHFSSADRYKYWREMAIQQAFVLKHLNVRCLLRR 714
T. cruzi           ASKLRALKTDSNEVLLKMDLESGHFSPSDRYRYWREMSFPQAFVLKHLNARTLLRR 714
                   **:** ****.****:****:*****.:***:* ** :: ********..* ***:
```