ENZYMATIC AND CRYSTALLISATION STUDIES OF CATL-LIKE TRYPANOSOMAL CYSTEINE PEPTIDASES

LAURELLE JACKSON
BSc (Hons) Biochemistry (*cum laude*)

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

Pietermaritzburg
2011

PREFACE

The experimental work described in this dissertation was carried out in the School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal, Pietermaritzburg, from January 2010 to November 2011, under the supervision of Professor 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.

_____

Laurelle Jackson (candidate)

_____

Professor Theresa H. T. Coetzer (supervisor)

DECLARATION – PLAGIARISM

I, Laurelle Jackson 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: ……

DEDICATION

I dedicate this work to my late grandfather (03/01/1929 – 26/11/2011), may your soul rest in peace but your legacy live on!

ABSTRACT

African animal trypanosomosis or nagana is a disease in livestock caused by various species of protozoan parasites belonging to the genus *Trypanosoma* particularly *T. congolense, T. vivax* and *T. b. brucei*. Nagana is the most important constraint to livestock and mixed crop-livestock farming in tropical Africa. Trypanosomes undergo part of their developmental life in their insect vector, the tsetse fly and part in their mammalian host.

Measures for eradicating the continent of the tsetse fly vector include insecticidal spraying, targeting and trapping. Vaccine development has been hampered by the generation of an inexhaustible collection of variant surface glycoproteins that trypanosomes possess and allow for evasion of the host immune system. Anti-disease vaccines aimed at reducing the symptoms of the disease rather than killing the parasite itself have been demonstrated as an alternative approach. Trypanotolerant cattle are able to protect themselves from the disease-associated symptoms. They are able to mount a better antibody response to the CATL-like cysteine peptidase, *Tco*CATL, compared to trypanosusceptible breeds. Bovine trypanosomosis, however, continues to be controlled primarily by trypanocidal compounds such as isometamidium chloride, homidium and diaminazene that have been developed more than 50 years ago and consequently drug resistance is widespread. Trypanosomal cysteine peptidases have also been proven to be effective targets for chemotherapeutics. *Tcr*CATL, inhibited by the vinyl sulfone pseudopeptide inhibitor K11777, was effective in curing or alleviating *T. cruzi* infection in preclinical proof-of-concept studies and has now entered formal preclinical drug development investigation.

Understanding enzymatic as well as structural characteristics of pathogenic peptidases is the first step towards successful control of the disease. To date no such characterisation of the major cysteine peptidases from *T. vivax* has been conducted. Although the major cysteine peptidase from *T. vivax*, *Tvi*CATL, has not been proven as a pathogenic factor yet, its high sequence identity with the pathogenic counterparts such as *Tcr*CATL and *Tco*CATL hold much speculation for *Tvi*CATLs role in pathogenocity.

In the present study, native *Tvi*CATL was isolated from *T. vivax* Y486, purified and characterised. *Tvi*CATL showed to have a general sensitivity to E-64 and cystatin and has a substrate specificity defined by the S_2 pocket. *Tvi*CATL exhibited no activity towards the

CATB-like substrate, Z-Arg-Arg-AMC but was able to hydrolyse Z-Phe-Arg-AMC, the CATL-like substrate. Leu was preferred in the P_2 position and basic and non-bulky hydrophobic residues were accepted in the P_1 and P_3 positions respectively. Similar findings were reported for *Tco*CATL. The substrate specificity of *Tvi*CATL and *Tco*CATL does argue for a more restricted specificity compared to *Tcr*CATL. This was based on the Glu333 in *Tcr*CATL substituted with Leu333 in *Tvi*CATL and *Tco*CATL. In the case of *Tcr*CATL, the Glu333 allows for the accommodation of Arg in the P_2 position. Like other trypanosomal cysteine peptidases, *Tvi*CATL was inhibited by both chloromethyl ketones, Z-Gly-Leu-Phe-CMK and H-D-Val-Phe-Lys-CMK. Determining further structural and functional characteristics as well as whether *Tvi*CATL, like the *T. congolense* homolog, *Tco*CATL, acts as a pathogenic factor, would be important information to the designing of specific chemotherapeutic agents.

To date, *Tcr*CATL and *Tbr*CATL (from *T. b. rhodesiense*) are the only trypanosomal CATL-like cysteine peptidases been crystallised and their tructures solved. This advantage has allowed for the directed design of synthetic peptidase inhibitors. The crystal structure of *Tco*CATL will be of major significance to the design of specific chemotherapeutic agents. Furtherrmore, understanding the dimeric conformation of *Tco*CATL is important for vaccine design as immune responses are likely to recognise the dimer specific epitopes. In the current study, the catalytic domain of *Tco*CATL and *Tvi*CATL, were recombinantly expressed in *Pichia pastoris* and purified to homogeneity. The *T. congolense* cysteine peptidase pyroglutamyl peptidase (PGP), also proven to be pathogenic in *T. b. brucei*, was recombinantly expressed in *E. coli* BL21 (DE3) cells and also purified to homogeneity. Purified cysteine peptidases along with previously purified *Tco*CATL dimerisation mutants, *Tco*CATL (H43W) and *Tco*CATL (K39F; E44P), possessing mutated residues involved in *Tco*CATL dimerisation, as well as the mutant proenzyme *Tco*CATL (C25A), were screened for crystallisation conditions using the Rigaku robotic crystallisation suite. One-dimensional needle-like crystals were found for *Tco*CATL (K39F; E44P). Optimisation of the *Tco*CATL (K39F; E44P) crystals were analysed for X-ray diffraction. The poor diffraction pattern prompted further optimisations for better crystal quality, which is presently underway. The crystal structure of *Tco*CATL, with some of the residues involved in dimerisation mutated, will be pivotal in understanding the dimerisation model. Furthermore, the information about the structure will be valuable for vaccine design and chemotherapeutics development.

ACKNOWLEDGEMENTS

I would like to express my gratitude and appreciation to the following people and institutions:

My supervisor, Professor Theresa Coetzer for all her assistance, support and time taken to help me learn the world of science and critically analyse my dissertation.

To Dr Greg Watson for all his knowledge and hours spent assisting me in my work.

To Charmaine Aherns and Robyn Hillebrand for all their assistance in financial and administrative matters.

To my fellow post-graduate students, Kayleen Brien, Richard Kangethe, Philia Vukea, Davita Pillay, Lauren Eyssen, Sabelo Hadebe, Melissa Govender, David Choveaux, Robert Krause, Jacky Viljoen, Kyle Goestch, Celia Snyman and Daniel Bohnen for all the happy times.

The National Research Foundation for the financial support.

To special friends that always make me laugh, Candice Wortmann, Cheralynn Miller and Sarah Choveaux.

To Ash and Mike, because we the three best friends that anybody could ever have!

To Gareth, for love, strength and sharing every happy memory with me.

To Mom and Dad for always being there and for your loving and caring nature.

To my grandparents for support, guidance and love.

To Holly, Forry, Megan and Emma for being the cutest pets and always making me smile.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

2x YT	2 × yeast extract, tryptone
ΔFLVp	full length inactive mutant of vivapain
A_x	absorbance at x nm
AEBSF	4-(2-aminoethyl)benzenesulfonylfluoride
AMC	7-amino-4-methylcoumarin
AMT	acetate-MES-Tris
BCA	bicinchoninic acid
Bis	N,N'-methylenebisacrylamide
Bis-Tris	2-bis(2-hydroxyethyl)amino-2-(hydroxymethyl)-1,3-propanediol
BMGY	buffered glycerol complex medium
BMM	buffered minimal medium
Boc	butyloxycarbonyl
BSA	bovine serum albumin
C-terminal	carboxy terminal
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E-64	L-*trans*-epoxysuccinyl-leucylamido(4-guanidino)butane
EDTA	ethylenediaminetetra-acetic acid
ELISA	enzyme-linked immunosorbent assay
g	relative centrifugal force
HRPO	horseradish peroxidase
[I]	inhibitor concentration
IAA	iodoacetic acid
IAM	iodoacetamide
IgG	immunoglobulin G
IgY	immunoglobulin Y

ISG	invariant surface glycoprotein
k_{ass}	rate of complex association
K_{av}	availability constant
k_{cat}	turnover number
kDa	kilo-Dalton
k_{diss}	rate of complex dissociation
K_i	inhibition constant
$K_{i(app)}$	apparent inhibition constant
K_m	Michaelis-Menten constant
k_{obs}	pseudo first-order inhibition rate constant
MEC	molecular exclusion chromatography
MeoSuc	methoxy-succinyl
MES	acetate-2(N-morpholino)ethanesulfonic acid
min	minute
N-terminal	amino terminal
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PSG	phosphate saline glucose
PEG	polyethylene glycol
PMSF	phenylmethylsulfonylfluoride
RT	room temperature (22 ± 2 ^{o}C)
[S]	substrate concentration
SDS	sodium dodecyl sulfate
Suc	succinyl
$t_{½}$	half-life
TBS	tris buffered saline
TEMED	N,N,N’,N’-tetramethyl ethylene diamine

TLCK	N-tosyl-L-lysyl chloromethylketone
TPCK	N-tosyl-L-phenylalanyl chloromethylketone
Tris	2-amino-2-(hydroxymethyl)-1,3-propandiol
VAT	variant antigen type
V_e	elution volume
V_{max}	maximum velocity
V_o	void volume
v_0	initial velocity
VSG	variant surface glycoprotein
V_t	total column volume
YNB	yeast nitrogen base
YP	yeast extract, peptone
YPD	yeast extract, peptone, dextrose
Z	benzyloxycarbonyl

CHAPTER 1

LITERATURE REVIEW

1.1 African Trypanosomosis

Human African trypanosomosis, endemic to intertropical Africa, also known as sleeping sickness, is represented by two main species *T. brucei rhodesiense* and *T. b. gambiense* (Lecaille et al., 2002). It is estimated that at least 50 million people are at risk of infection with trypanosomes (http://who.int/tdr/diseases/tryp/default.htm). African animal trypanosomosis (AAT), caused by *Trypanosoma congolense*, *T. vivax*, *T. evansi* and *T. b. brucei*, is the most important constraint to livestock and mixed crop-livestock farming in tropical Africa. An area of 8.7 million km^2 of the continent is infested with the tsetse fly vector, *Glossina spp*, resulting in at least 46 million cattle and 70 million small ruminants being exposed to the risk of contracting the disease (Fig 1.1) (Kristjanson et al., 1999).

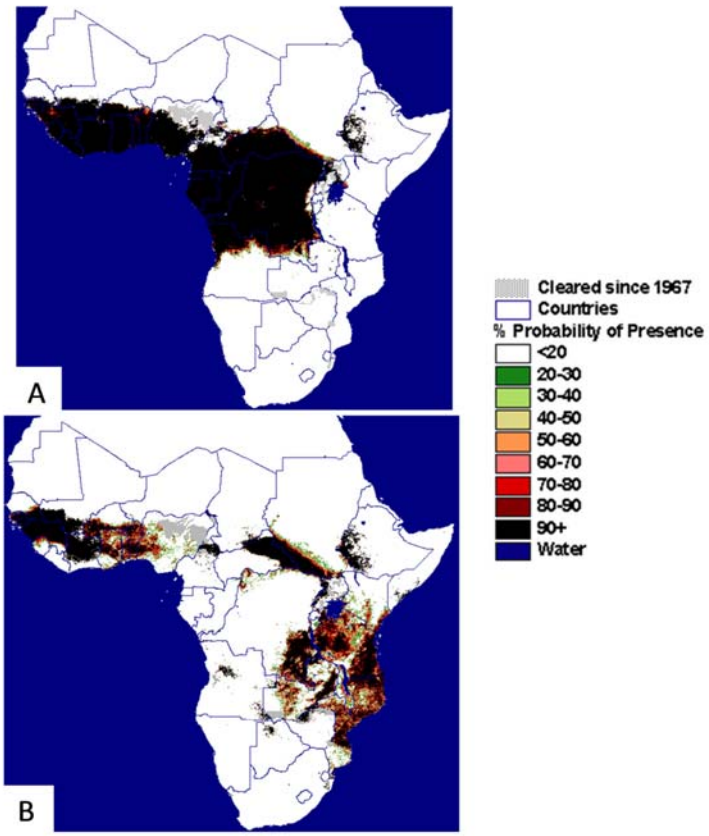


Figure 1.1 Distribution of (A) *Glossina palpalis* and (B) *Glossina morsitans* groups of tsetse flies predicted using satellite imagery (Wint and Rogers, 2000).

This disease affects 40 African countries and is costing the continent US $4.75 billion per year (Van den Bossche et al., 2010). The disease has caused considerable direct economical losses as costs are mainly attributed to those involved in diagnosis and treatment as well as the decrease in milk and meat production. The African continent produces one-seventieth of the animal protein per hectare that Europe produces (Kristjanson et al., 1999). This is generally a result of farmers' responses to the perceived risk of the disease, including the reduction and, in some cases, the exclusion of livestock from tsetse-infested grazing lands and reduced crop production due to insufficient animal draft power leading to further indirect economic losses. Furthermore, the International Livestock Research Institute (ILRI) stated that the impact of trypanosomosis is even greater than these figures suggest because many of the areas inhabited by tsetse flies are potentially the most agriculturally productive in Africa (http://www.ilri.org/InfoServ/Webpub/fulldocs/Ilrad90/Trypano.htm). There have been several approaches to controlling the disease such a trypanocides, controlling tsetse populations, searching for components for a multi-component vaccine and the use of livestock breeds that tolerate the disease (trypanotolerant breeds). The potential benefits of improved trypanosomosis control, in terms of meat and milk production alone, amount 700 million US dollars per annum in Africa (Kristjanson et al., 1999).

1.1.1 Classification

African animal trypanosomes belong to the order Kinetoplastida that comprises a group of protozoa defined by the presence of a characteristic organelle, the kinetoplast (Vickerman, 1969). This organelle is a large modified mitochondrion. Originally established classification systems of the parasite were based largely on the morphological and lifecycle characters in the absence of molecular data. The Kinetoplastida order was thus originally subdivided into the suborders Bodonina and Trypanosomatina (Vickerman, 1969). The parasitic bodonidae were represented by ectoparasites of fish and endoparasites of fish and snails. The second group comprised a single family of obligate parasites, Trypanosomatidae. There have been improvements in sampling taxa for phylogenetic studies using molecular techniques (Hamilton et al., 2004; Simpson et al., 2002; Stevens and Brisse, 2004). Based on new molecular data, a new system was generated which divided kinetoplasts into Prokinetoplastina (*Ichthyobodo* and *Perkinsiella*) and Metakinetoplastina (other bodonids and trypanosomatids) (Fig. 1.2) (Moreira et al., 2004).

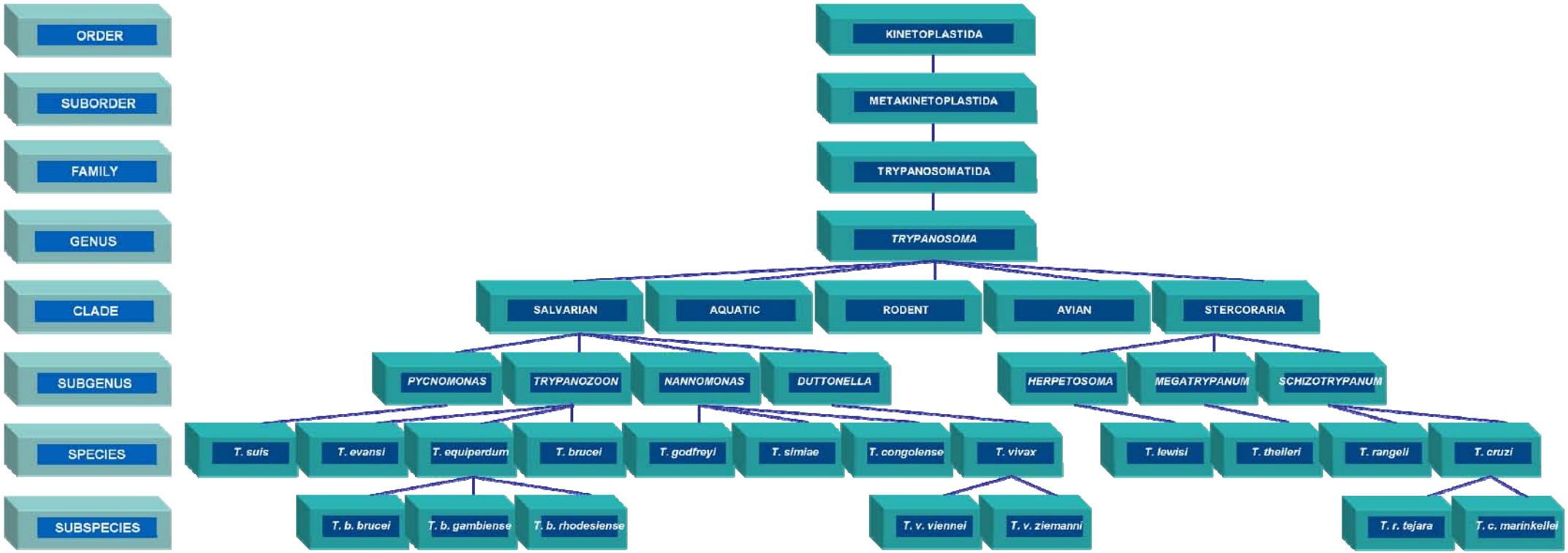


Figure 1.2 Classification of trypanosomes (Moreira et al., 2004; Stevens and Brisse, 2004).

Trypanosomatids have a digenetic lifestyle, they are carried in invertebrate vectors such as tsetse flies, sandflies and triatomine bugs and infect all classes of vertebrates (Stevens, 2008). The *Trypanosoma* genus comprises the four clades including Salivarian and Stercoraria clades that contain species of medical and veterinary importance and have become the focus of much research. The Stercoraria clade contains the species *T. cruzi*, which has a wide vertebrate host range and is the causative agent of Chagas' disease. *T. rangeli* is infective to humans whereas *T. thieleri* has a wide host range, with the major vector being tabanid flies. The only species from the subgenus *Duttonella* within the Salivarian clade, *T. vivax*, represents early stage adaptation to transmission by tsetse flies (Lukes et al., 1997) as they are unable to undergo cyclical development in tsetse flies. The *Trypanozoon* subgenus umbrellas the widely researched *T. brucei*. The subspecies *T. b. rhodesiense* and *T. b. gambiense* are the causative agents for human African sleeping sickness (HAT), both transmitted by tsetse flies. *T. congolense*, *T. simiae* and *T.godfreyi* are species of the subgenus *Nannomonas*. *T. congolense* along with *T. vivax* and *T. b. brucei* are frequently encountered pathogens of African livestock, causing the disease Nagana. On the other hand, *T. evansi*, part of the *Trypanozoon* subgenus, causes Surra, a severe wasting disease in livestock. *T. evansi* is mechanically transmitted by biting flies, namely tabanids.

1.1.2 Morphology of African trypanosomes

Trypanosomal parasites are unicellular organisms which vary between 8 and 24 μm. The streamlined parasites are tapered at both the anterior and posterior ends to assist in the efficient movement through the bloodstream. Trypanosomes have a highly organised exoskeleton (Vickerman, 1969) and along with the pellicular microtubules, which cover the entire cell surface beneath the cell membrane of the parasite, give the structure of the trypanosome as well as allowing movement and flexibility (Stevens and Brisse, 2004; Taylor and Authié, 2004). The flagellum, produced by the parabasal body (Fig.1.3) is essential for mobility. Unlike *T. vivax* and *T. brucei*, *T. congolense* lacks a free flagellum. This difference has been utilised for diagnostics using microscopy (Eisler et al., 2004). The invagination of the plasma membrane where the flagellum attaches to the trypanosome is known as the flagellar pocket. The flagellar pocket is generally the site for endo- and exocytosis, where protein trafficking and recycling of cellular proteins takes place.

The single mitochondrion that trypanosomes possess runs from the posterior to the anterior end. The posterior end of the mitochondrion comprises the specialised area known as the kinetoplast. The kinetoplast contains between 10-20% of the trypanosomal DNA (Vickerman, 1969). The kinetoplast is significantly larger in *T. vivax* compared to *T. congolense* which is exploited for diagnostics (Eisler et al., 2004).

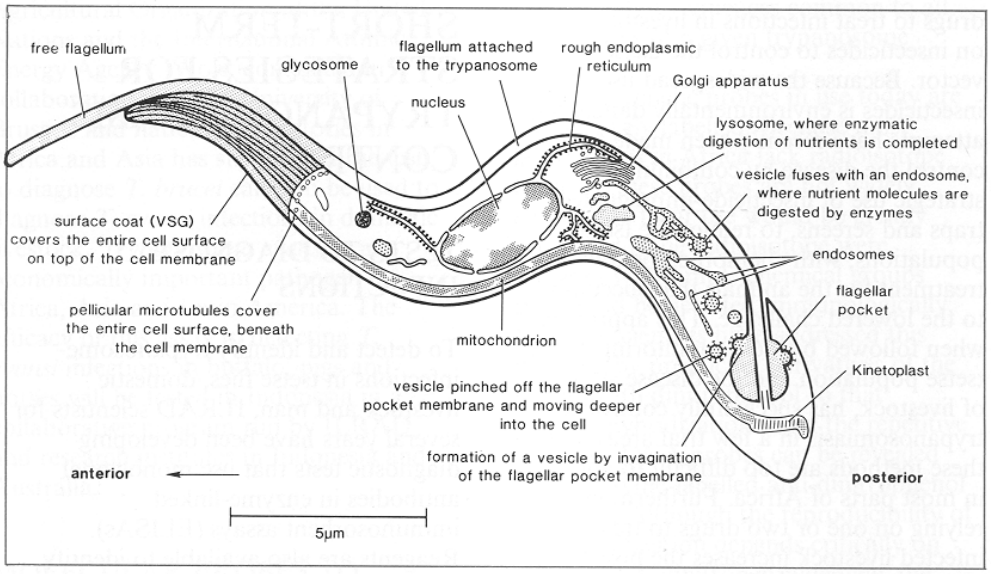


Figure 1.3 Schematic diagram of a trypanosome of the *T. brucei* group in its intermediate bloodstream form, illustrating the major organelles. (Acquired from International Livestock Research Institute's website www.ilri.org/InfoServ/Webpub/fulldocs/Ilrad88/Trypanosomiasis.htm accessed 02.10.11).

1.1.3 Life cycle of African trypanosomes

Trypanosomes are digenetic parasites and as a result of this the parasite undergoes a series of pleomorphic events that are necessary for the successful transmission from the vector to the host and *vice versa* (Fig. 1.4). Trypanosomes undergo part of their developmental life in their insect vector, the tsetse fly or other invertebrate depending on trypanosome species and part in their mammalian host (Roditi and Lehane, 2008). Three developmental forms of the trypanosome undergo cell multiplication: the non-infective procyclic and epimastigote forms in the tsetse fly and the long slender bloodstream forms in the mammalian host (El-Sayed et al., 2000). The developmental forms that do not undergo multiplication are the infective matacyclic and short stumpy (not shown in Fig. 1.4) forms.

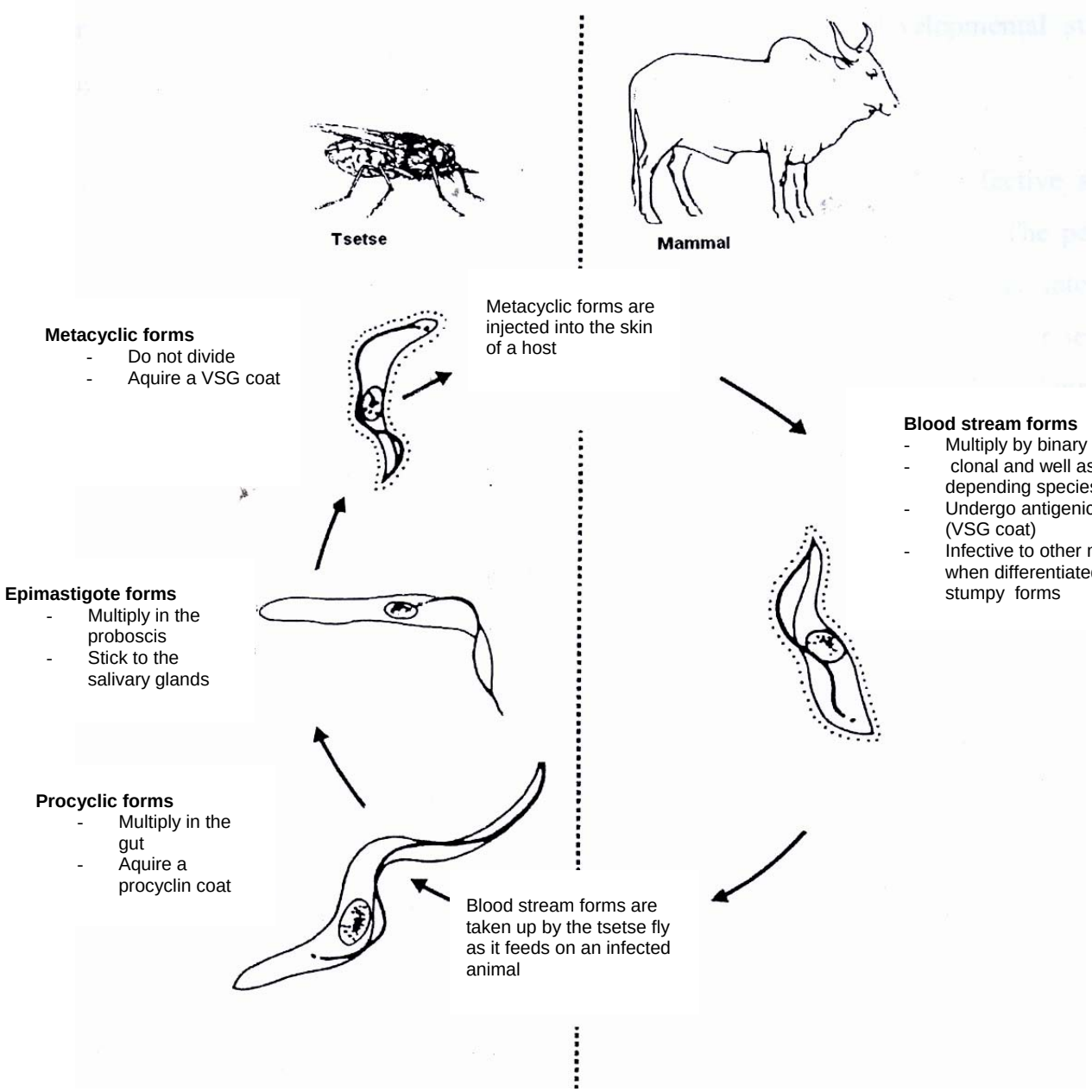


Figure 1.4 The life-cycle of *T. congolense* (Authié, 1994).

As the fly takes in a blood meal from the infected mammalian host, there are two forms that can potentially be ingested. These include the normal slender bloodstream forms or the short stumpy, non-replicating forms. Only the short stumpy forms of the trypanosome are thought able to survive the onslaught of potent peptidases present in the saliva and midgut of the tsetse fly. The variable surface glycoprotein (VSG) coat of the surviving stumpy forms is shed and replaced with an invariant glycoprotein coat composed of procyclic acidic repetitive proteins (PARPs), commonly referred to as procyclins, which are attached

to the membrane via glycosylphosphatidylinositol (GPI) anchors (El-Sayed et al., 2000). The procyclic forms colonise the ectoperitrophic space and reach very high densities in the midgut. To complete part of the life-cycle in the insect vector, the procyclics forms must migrate to the salivary glands via the foregut and the proboscis. In the case of *T. vivax*, development within the vector is confined entirely to the proboscis and lacks the procyclic stage (Stevens and Brisse, 2004)

In the salivary glands, the epimastigote form develops. This form attaches to the epithelium and proliferates to give rise to the mammalian infective non-dividing metacyclic form. It is the metacyclic form that expresses the VSG that is evident on the surface of the parasite. These metacyclic forms are transmitted to a susceptible mammal during the vector's blood meal. The existence of mating in trypanosomes has been documented for *T. brucei* (Gibson and Stevens, 1999) and *T. congolense* (Holzmuller et al., 2010). As a result of this, trypanosomes can use strategies from clonal reproduction, by binary fission, to sexual recombination by mating, to adapt to various environmental conditions.

1.1.4 Genetic organisation of trypanosomes

Trypanosomes possess two genomes, one within the nucleus and the other within the kinetoplast. The kinetoplast DNA is arranged into an interlinked network of minicircles (1 kb) and maxicircles (~ 22 kb) (El-Sayed et al., 2000). The maxicircles have conserved sequences and encode for mostly ribosomal RNA and mitochondrial proteins whilst minicircles occur in a very high copy number and comprise variable sequences that encode for a wide variety of proteins (Vather, 2010). To date, the *T. brucei* nuclear genome has been completely sequenced (Berriman et al., 2005). The knowledge of nuclear genome has provided information for trypanosomal genomes such as *T. congolense* and *T. vivax* (Morty et al., 2006). *T. brucei* has a 3.5×10^7 bp genome which is approximately the size of the *Plasmodium falciparum* genome (Dame et al., 1996). The nuclear genome of trypanosomes is of low complexity, with most of it being diploid. The VSG genes as well as the VSG expression sites and some of the mini-chromosomes are haploid (El-Sayed et al., 2000).

1.1.5 Antigenic variation in trypanosomes

Antigenic variation involves changes in the identity of the VSG (Vickerman, 1978). The entire surface of the bloodstream and metacyclic forms of the trypanosome, including the

flagellum, is covered with VSG. Each variant is termed a distinct variable antigen type (VAT). It is thought that the surface coat is there to provide general protection against host resistant mechanisms and therefore shield invariant surface antigens from immune recognition. Furthermore, the antigenic variation system in trypanosomes ensures survival within the mammalian bloodstream and allows transmission to the next host via the tsetse vector (Morrison et al., 2009).

The stochastic switching to a different VAT which allows for evasion of the host immune system causes prolonged trypanosome infection (Barrett et al., 2003; Donelson, 2003; McCulloch, 2004). The rate of VSG switching is up to 10^{-3} switches/parasite/5-10 h generation time (Turner and Barry, 1989). VSG switching in itself is a mechanism to evade the host immune system. However, VSG also has several other effects on the immune system such as induction of autoantibodies by the non-specific polyclonal activation of B-cells (Buza and Naessens, 1999) as well as cytokines, namely tumour necrosis factor (TNF)-α (Morrison et al., 2009).

Sites for VSG transcription (VSG expression site) are found directly proximal to the telomeres (Hertz-Fowler et al., 2008). Subtelomeres also harbour silent VSG arrays (Berriman et al., 2005). The VSG gene array contains a large portion of pseudogenes which combine to form functional VSG ‘mosaics’ (Marcello and Barry, 2007), which provide a critical resource during antigenic variation. In order for VSGs to switch their identity certain switching mechanisms have to occur. There are two modes of switching: transcriptional switching, where a single actively transcribed VSG expression site is silenced and a silent expression site becomes actively transcribed, and recombination reactions (Morrison et al., 2009).

1.2 Pathogenesis of African animal trypanosomosis

There is a large amount of intraspecies variation in the pathogenocity of different parasite stocks, especially stocks isolated form distinct geographical regions (Taylor and Authié, 2004). Haematic trypanosomes such as *T. congolense* and *T. vivax* are found in the blood vessels of hosts, whilst humoral trypanosomes, such as *T. b. brucei*, *T. evansi* and *T. equiperdum* are found in the extravascular compartment and tissues (Lonsdale-Eccles and Grab, 2002). Furthermore, acute haemorrhagic and non-haemorrhagic disease can be found

in cattle when infected with *T. vivax* from East and West Africa respectively. Generally, there are three main phases of bovine trypanosomosis. These include the early acute phase, stabilisation and chronic phase (Taylor and Authié, 2004). In the acute phase, there is a continuous detectable level of trypanosomes in the blood. Fever is high and anaemia can develop (Andrianarivo et al., 1995; Murray and Dexter, 1988). Furthermore the spleen and lymph nodes are enlarged. Very often death of infected animals can occur in this acute phase as there is a continuous drop in the haematocrit value (packed cell volume) (Taylor and Authié, 2004). Stabilisation only occurs in cattle that are able to stabilise the packed cell volume. Generally stabilisation can occur after 6-8 weeks, whilst the parasitaemia becomes intermittent. The packed cell volume may fluctuate at low levels for months or years. The chronic stage occurs when cattle develop cachexia, impaired milk productivity and oestrous cycle. The chronic phase can lead to death and self-cure to, a persistent carrier state (Taylor and Authié, 2004).

During the infection process, once drained from the afferent and efferent lymphatics, the lymphoid organs, liver, bone marrow, heart, central nervous system, endocrine and reproductive organs become affected (Mattioli et al., 1999). The metacyclic trypanosomes deposited in the skin undergo multiplication. Extensive oedema occurs and a chancre develops as a result of degranulation of mast cells in the skin. After a period of multiplication trypanosomes appear in the afferent lymph nodes. The cellularity of the afferent lymph increases significantly and is characterised by $CD4^+$, $CD8^+$ and B lymphocytes (Sileghem and Naessens, 1995; Uzonna et al., 1998). Splenomegaly is a feature of the ‘acute’ or parasitaemic phase of the infection and is mainly a result of lymphocyte sequestration and an expanded macrophage population (Noël et al., 2002). This is notably seen in the red pulp of the spleen, as there is an expansion of the red pulp.

The co-evolution of trypanosomes with their host have allowed for trypanosomes to divert and alter the numerous steps that lead to the hosts immune response. As a result of this the hosts’ immune response does not necessarily lead to protection but to immunopathology disorders. Immunopathology disorders include production of autoantibodies, immune suppression of T- and B-lymphocytes and dysregulation of the cytokine network (Vincendeau and Bouteille, 2006). The pluripotent cytokine, IL-10, is upregulated in mononuclear cells derived from the peripheral blood, lymph nodes and spleen of infected cattle. IL-10 has the ability to down-regulate IFN-γ. In terms of Type-1 cytokine

production, susceptible mice produce more IFN-γ than resistant mice (Tabel et al., 2000) and inhibiting IFN-γ causes a marked increase in survival (Uzonna et al., 1998). It has been proposed that mice die from Systemic Inflammatory Syndrome (SIRS) mediated by IFN-γ (Stijlemans et al., 2007).

1.3 Diagnosis of African animal trypanosomosis

Generally herdsman or technical personnel conducting diagnostic tests operate under the most adverse physical and economic conditions. African herdsmen usually rely on the clinical features of the disease for diagnosis The clinical signs of acute bovine trypanosomosis include anaemia, weight-loss, roughness of hair coat, enlargement of peripheral lymph nodes, pyrexia and abortion (Eisler et al., 2004). The presumptive diagnosis, based on clinical signs, is usually followed by the administration of trypanocidal drugs, with a positive response to treatment being interpreted as a confirmation of diagnosis.

Parasitological diagnosis rely on blood smears fixed with methanol and stained with Giemsa followed by light microscope examination (Eisler et al., 2004). This technique is not sensitive and cannot detect low parasite levels which are often characteristic of infection in livestock. The haematocrit centrifugation technique (Woo, 1970), involves centrifugation of a microhaematocrit capillary tube containing the infected blood sample and examination of the buffy coat-plasma interface. In an improvement of this method the buffy coat plasma interface is placed on a microscope slide and dark-ground or phase contrast illumination is used to detect the parasites (Murray et al., 1977). Other methods for diagnosis were developed such as the sub-inoculation method, where blood is taken from a suspected infected case and injecting it into another species such as immunosuppressed rodents or mice where levels of parasitaemia develop very quickly (Eisler et al., 2004). However, like in the case of *T. vivax*, mouse sub-inoculations fail to pick up infections. An *in vitro* culture method has also been developed. The kit for the *in vitro* isolation of trypanosomes (KIVI) was developed for *T. b. gambiense* in human trypanosomosis infections (Aerts et al., 1992) and has also been shown to detect *T. b. brucei* in animals with more sensitivity than parasitological techniques (McNamara et al., 1995).

The low levels of parasitaemia has made serodiagnostics appropriate for the detection of trypanosomes in animals. The card agglutination trypanosomosis test (CATT) (Akol et al.,

1999) is very suitable for field use in Africa, whereby infected blood serum antibodies will agglutinate with card antigens (antigens blotted onto a nitrocellulose card) based on the antigenic VATs. This method is suitable for *T. b. gambiense* which has highly conserved VATs, however, for *T. congolense* and *T. vivax,* there is difficulty in identifying suitable VATs and therefore this assay is not suitable (Luckins, 1992). The CATT/*T. evansi* diagnostic kit uses a freeze dried suspension of purified, fixed and stained trypanosomes of a cloned predominant VAT (VAT RoTat 1.2) of *T. evansi* on a card agglutination to detect antibodies from *T. evansi* infected animals (Reid and Copeman, 2003).

An antibody-detection ELISA relies on the specific antibodies in the infected sample to bind to a specific antigen and is itself detected by means of an enzyme-labelled anti-species (anti-bovine) conjugate (Eisler et al., 2004). The major advantage of any ELISA detection is that it can be performed without very specialised equipment. Crude trypanosomal antigen preparations fractionated at different percentages of ammonium sulfate have been used for the detection of antibodies in *T. evansi* infected cattle (Reid and Copeman, 2003). The three antigen preparations, labelled according to the ammonium sulfate fractionation, 20-40%, 40-50% and >50%, were compared against the CATT/*T. evansi* test (Bajyana Songa and Hamers, 1988). The “40-50%” antigen preparation had the greatest discriminatory power between non-infected and *T. evansi* infected cattle. Furthermore, the antigen preparation was significantly more sensitive than the CATT/*T. evansi* test. The indirect ELISAs using denatured crude trypanosomal antigen preparations for *T. congolense* (TcAGd) and *T. vivax* (TcAGd), which resulted in the development of an ELISA kit (I-TAB ELISA) has been used for the detection of anti-trypanosomal antibodies in bovine serum (Rebeski et al., 2001). This ELISA has also been used for *T. viax* and *T. congolense* infected goats (Lejon et al., 2003). The ELISA was able to detect antibodies from *T. congolense* infected goats with 100% sensitivity. However there was extensive cross reactions of the ELISA with sera from *T. vivax* and *T. brucei* infected goats. One of the most important obstacles to antibody-detection ELISAs is that antibodies can persist in the host far longer than the infectious agent does (Boulangé et al., 2002; Van den Bossche et al., 2000). However, the heat shock protein 70/BiP inhibition ELISA for the diagnosis of bovine trypanosomosis was able to show seroconversion within three months after drug treatment in experimentally infected animals, suggesting that it could be a useful tool for monitoring the efficiency of treatment (Bossard et al., 2010). Furthermore, the antigen-detection ELISA overcomes the obstacle by detecting antigen rather than antibody in the

infected host. The double-sandwich ELISA method was shown to detect antigens, with antisera raised against *T. congolense*, in rabbits and goats within 10-14 days of infection in *T. congolense* and these antigens disappeared 21 days after trypanocidal drug treatment (Rae and Luckins, 1984). An antigen detection ELISA for species specific diagnosis, developed at the Animal Production of Health unit, Atomic Energy Agency (IAEA) Agriculture laboratory, Seibersdorf, Austria, uses the monoclonal antibodies [T.b.r.7 (IgM)] (Nantulya et al., 1987), [T.c.39 (IgM)] (Masake and Nantulya, 1991), [T.v.27 (IgG_1)] (Nantulya et al., 1992) for *T. b. brucei*, *T. congolense* and *T. vivax* respectively. However, this antigen detection ELISA does have poor sensitivity and specificity (Eisler et al., 1998).

Molecular methods, such as PCR, for detecting livestock-infective trypanosomes has evolved as one of the most specific and sensitive methods of diagnosis. Some research has focused on multiple species identification with a single primer (Desquesnes et al., 2001). Furthermore, a PCR-ELISA assay, using universal PCR amplifying trypanosome DNA followed by an ELISA-based hybridization with three highly-specifc probes, has been developed and can differentiate between *T. congolense* and *T. vivax* (Cabrera et al., 2009). In spite of the good specificity and sensitivity, molecular biology techniques are not commonly used due to cost and lack of skilled personnel in diagnostic laboratories in developing countries (Thekisoe et al., 2007). The loop-mediated isothermal amplification (LAMP) of *T. b. gambiense*, *T. b. rhodesiense*, *T. congolense*, *T. cruzi* and *T. evansi* DNA with specific primers overcomes many of the molecular diagnostic drawbacks (Thekisoe et al., 2007). Under isothermal or constant temperature conditions, auto-cycling strand displacement of trypanosomal DNA is performed by the *Bst* DNA polymerase (Thekisoe et al., 2007). The amplification is rapid, taking about a 1 h, requires only a heat bath, can amplify trypanosome DNA from blotted blood (Mori et al., 2001) on filter paper and produces large amounts of DNA that can visualised with the naked eye in the test tube (Mori et al., 2001).

1.4 Trypanotolerance

Many years of selection in tsetse-infested areas allowed for taurine breeds of cattle in West-Africa to develop a reduced susceptibility to trypanosomosis. The term trypanotolerance was defined as the trait that confers the capacity to survive and remain productive after trypanosome infection (Murray et al., 1982). Generally, trypanotolerant

cattle have reduced mortality, blood trypanosome levels and severe anaemia and their reproductive performance is better than indicine breeds such as Boran cattle (*Bos indicus*) (Naessens, 2006). Such observations were made under laboratory conditions between trypanotolerant N'Dama calves (*B. taurus*) and susceptible Boran calves. Trypanotolerance has been suggested to be multi-factorial, involving both non-immunological and immunological mechanisms (Authié, 1994).

Efforts to identify genes that may confer trypanotolerance have been made. Serial analysis of gene expression (SAGE) has been used to identify the change of leukocyte number in N'Dama after infection by *T. congolense* (Maillard et al., 2004; Maillard et al., 2005). N'Damas have a comparative advantage over Boran cattle in erythropoietic responsiveness. The receptor for the glycoprotein hormone controlling erythropoiesis, erythropoietin, has significantly higher transcription levels in the bone marrow of N'Damas, compared to Borans with *T. congolense* infection.

The role of hematopoietic cells in trypanotolerance has also been previously investigated. As illustrated by placing a Boran and an N'Dama embryo in the recipient mother, resulting in bone marrow chimerism in twin calves, N'Dama cattle haemopoietic cells in anaemia control and parasitaemia were not linked (Naessens et al., 2003). It has also been recently shown that expression levels of serum amyloid A, an acute phase protein, working as part of the host's effector molecules in the innate immune system is elevated in N'Dama cattle when compared to Boran cattle in experimental infection with *T. congolense* (Meade et al., 2009)

There is no difference between trypanotolerant and susceptible cattle regarding the production of antibodies against the main target for the humoral immune response, VSG (Pinder et al., 1984). There are however, parasitic antigens that are released into the bloodstream of the host upon parasite lysis by anti-VSG antibodies. These antigens can induce pathology and trypanotolerant cattle may possess the ability to resist the disease better through a more effective immune response to these antigens (Authié, 1994). Western blotting has illustrated that a 33 kDa CATL-like cysteine peptidase, namely *Tco*CATL (Mbawa et al., 1992), induce the production of reactive IgG1 antibodies in trypanotolerant cattle (Authié et al., 1993). This antibody response is one of the clear differences between N'Dama and Boran cattle (Taylor et al., 1996).

1.5 Control of African Animal Trypanosomosis

1.5.1 .Chemotherapy

Over most of sub-Saharan Africa, bovine trypanosomosis continues to be controlled primarily by chemotherapeutic agents such as Samorin® (isometamidium chloride), Ethidium® (homidium) and Berenil® (diaminazene). However these main drugs registered to treat African trypanosomosis have mostly been used for more than 50 years (Kroubi et al., 2011). Ongoing research is, however, facilitating the development of drugs that are more therapeutic, combat resistance and possess fewer side effects.

The compound isometamidium chloride (ISM; Samorin®, Trypanidium®, Veridium®) (Table 1.1) has been used a prophylactic drug for livestock suffering primarily from *T. congolense* infection and can provide up to six months protection. ISM is a conjugate of homidium (Ethidium®) and diaminazene (Berenil®) molecules (Delespaux and de Koning, 2007). The main mode of action of ISM is proposed to be the cleavage of kinetoplast DNA-topoisomerase complexes, causing the desegregation of the minicircle network within the kinetoplast (Shapiro and Englund, 1990). The kinetoplast is the primary site for ISM accumulation in *T. brucei* and *T. congolense* (Boibessot et al., 2002).

Diaminazene has limited prophylactic properties and major therapeutic properties. Diaminazene is relatively cheap and widely available (Bruning, 1996). This is possibly why it is used as a first-line of treatment for sick livestock. However rapid clearance of the drug from the body has contributed to the development of resistance (Delespaux and de Koning, 2007).

There are to two major strategies for trypanosomal drug usage in African trypanosomosis infections. The first is routine block treatments where prophylactic drugs, mainly ISM, are used at predetermined intervals based on prophylaxis duration. The second is strategic block treatments using prophylactic drugs when animals succumbing to infection reach a predetermined threshold of parasitaemia (Holmes et al., 2004). Recent work has focused on drug delivery systems to improve drug therapeutic efficacy, optimise the drug's pharmacokinetic properties and reduce its toxicity (Kroubi et al., 2011). The ultimate goal for drug delivery systems is to optimise ADME (absorption, distribution, metabolism and elimination) and at present several strategies aimed at improving ADME are under

development which include the inclusion of drugs into liposomes, polymeric miscelles, cyclodextrins and nanoparticles (Kroubi et al., 2011). Diminazene liposomes have been studied in a mouse model of *T. b. brucei* and *T. evansi* infection (Yongsheng et al., 1996). The experiment demonstrated that encapsulated diminazene was more active than the free drug, did not give side effects and had a longer half-life and efficacy in the mice. A further study showed that a lipid, water-insoluble drug conjugate made of distearate, tristearate and diolate with diminazene has increased adherence to the brain endothelial vessels, thus being able to cross the blood-brain barrier (Chang et al., 2009).

1.5.2 Tsetse control

One of the principal strategies in controlling African trypanosomosis is to eradicate or control tsetses fly populations. Due to the absence of vaccines and effective and affordable drugs, the African Union, in 2002, initiated a new continent-wide initiative to rid the entire continent of the tsetse vector: the Pan African Tsetse and Trypanosomosis Eradication Campaign (PATTEC) (Symeonakis et al., 2007). The three major strategies that have been adopted to rid the continent of such a vector include insecticidal spraying, insecticidal traps and the sterile insect technique. Targets and traps have been used extensively in small agriculture settings (Vale, 1993), focusing on trapping of tsetse flies on water (Laveissiere et al., 2011).

Tsetse flies are susceptible to most insecticides. However, as part of the life cycle, the juvenile tsetse fly development takes place several centimetres below ground and this determined the approach to insecticide delivery such as ground spraying using hand-held or vehicle mounted fogging machines as well as sequential aerosol technique using aircraft (Allsopp and Hursey, 2004). Insecticides include the organochlorines such as DDT (Campbell and West, 1952), dieldrin and endosulfan and the more recently developed synthetic pyrethroids (Allsopp and Hursey, 2004). Bioaccumulation of the organochlorines made them very effective over a long time period but ultimately led to their international ban (Allsopp and Hursey, 2004). Endosulfan does not have the same potential for bioaccumulation and because of its wide availability at reasonable cost, was a suitable choice for non-residual aerosol applications from aircraft.

However, the discovery of synthetic pyrethroids led to the development of deltamethrin that is much more powerful than the organochlorines and has low mammalian toxicity and a very broad invertebrate spectrum (Elliott et al., 1974). Sterile insect technique is a

genetic population suppression approach that involves sustained, systematic release of irradiated sterile male insects among the wild population (Aksoy, 2003). In combination with trapping and insecticide spraying, it has been a successful tsetse eradication technique in contained areas such as Zanzibar (Vreysen et al., 2000).

1.5.3 Vaccine development

As trypanosomes are able to continuously evade the host antibody response with the generation of an inexhaustible collection of VSGs, using killed parasites as a vaccine is not an option. Non-variable antigen-based vaccines have been explored. Trypanosomes do express a number of less variable surface antigens. These include the multiple copy-number invariant surface glycoproteins (ISGs), evenly distributed on the surface of the bloodstream-stage trypanosomes. Induction of antibodies against ISG75 was a prominent result during flagellar pocket vaccination, indicating that this protein is immunogenic by itself, despite its heavy glycosylation (Radwanska et al., 2000). However mice immunised with ISG75 demonstrated no protection upon challenge with *T. b. brucei* (Tran et al., 2008). It was found that the effective B cell memory response was abolished. As Magez et al. (2009) observed that vaccination still remains the best theoretical option, in order for an anti-trypanosome vaccine to be effective, it should have the ability to eliminate all circulating parasites prior to B cell memory suppression or destruction.

Alternatively, transmission blocking vaccines (TBVs) are aimed at reducing transmission of the trypanosome through immunisation of the host against insect parasite stages. The antibodies circulating in the host blood would be transferred to the tsetse fly in the blood-meal. This type of vaccine would ultimately interfere with the parasite life cycle in the tsetse fly by targeting specific interactions that are required for parasite development, reduce the vector fitness and block parasitaemia onset in the host (Magez et al., 2009). The presence of anti-procyclic *T. b. brucei* and *T. congolense* antibodies in tsetse blood-meals

Table 1.1 Trypanocidal drugs for the treatment of African Animal Trypanosomosis[†]

Pharmaceutical class ***Drug name***	**Pharmaceutical characteristics**	**Activity against**	**Drawbacks**
Diamidine *Diaminazene** (Berenil®, Veriben®) 1955	Intramuscular and intravenous. Uptake by carrier-mediated transport (Bray et al., 2003; de Koning et al., 2004). Inhibition of the kinetoplasmatic DNA biosynthesis (Balana-Fouce et al., 1986; Karvonen et al., 1985).	*T. congolense* *T. vivax* *T. brucei* *T. evansi*	Licensed just for veterinary use. Does not cross the blood-brain barrier (Reinert, 1999). Rapid decrease of the drug plasma concentration and slow elimination. Side effects and numerous resistant strains (Ainanshe et al., 1992; de Koning et al., 2004; Sinyangwe et al., 2004).
Melaminophenyl arsenical *Melarsomine* (Cymelarsan®) 1985	Intravenous, intramuscular. Especially effective on domestic animals infected by *T. evansi* (Lun et al., 1991) and *T. equiperdum*. Not effective against *T. congolense* and *T. vivax.*	*T. evansi* *T. equiperdum*	Resistant strains. Rapidly Metabolised in the plasma and major side effects.
Phenanthridine *Homidium*®‡ (Ethidium®), 1952	Intravenous, intramuscular. Uptake by passive diffusion. Inhibition of the DNA biosynthesis.	*T. congolense* *T. vivax* *T. brucei*	Resistant strains and numerous side effects (Boid et al., 1989; Ndoutamia et al., 1993).
Isometamidium‡ (Samorin®/Veridium®), 1960		*T. evansi*	

† Table adapted from Kroubi et al. (2011) and Holmes et al. (2004)

* Group I compounds: produce their *in vivo* trypanocidal effect immediately

‡ Group II compounds: produce their *in vivo* trypanocidal effect only after a latent period of 24 hours.

have been reported to suppress the development of the respective parasites in *Glossina morsitans* (Maudlin et al., 1984). However, this vaccine approach does have an ethical problem in that it does not protect the vaccinated individual against the infection but would however ultimately result in reduced numbers of infectious vectors.

Anti-disease vaccines, aimed at reducing the symptoms of the disease rather than killing the parasite, first developed for malaria (Playfair et al., 1990), have been considered as an alternative approach to conventional vaccines (Authié, 1994). Trypanotolerant cattle do not show the disease-associated symptoms (Authié et al., 2001) and are able to mount an antibody response to the proteolytic enzyme, *Tco*CATL. The artificial induction of antibodies against the peptidase could render a diseases-susceptible host more tolerant (Authié et al., 1992; Lalmanach et al., 2002).

1.6 Proteolytic enzymes

Proteolytic enzymes release amino acids, peptides, and proteins from larger peptides and proteins (Rawlings et al., 2008). Peptidases, often synonymous with proteolytic enzymes, catalyse the cleavage of amide linkages or peptide bonds in proteins (Rawlings et al., 2008) and can either have endopeptidase activity whereby the polypeptide chain is cleaved within the chain or amino or carboxy peptidase activity where only the N- or C- terminal residue is cleaved off (Sajid and McKerrow, 2002). There are seven classes of peptidases, four catalytic types, cysteine, serine, aspartic and metallo, were recognised by Rawlings and Barrett (1993). The fifth catalytic type, possessing an N-terminal threonine that acts as a nucleophile, was later discovered (Rawlings and Barrett, 1999). Glutamate (Fujinaga et al., 2004) and the more recently discovered peptidases that cleave themselves at asparagine residues (Rawlings et al., 2011) make up the sixth and seven classes respectively. The clans, within the seven classes, usually show evolutionary relationships by their similar tertiary structures, catalytic site residues or sequence motifs (Caffrey and Steverding, 2009).

1.6.1 Cysteine peptidases

Cysteine peptidases are also referred to as thiol or sulfhydryl peptidases and contain a cysteine in the catalytic site. Cysteine peptidases are a diverse group of enzymes found in all living organisms. In humans they are important in the immune system, protein renewal processes and the resorption of bone cartilage, where imbalances of these peptidases can

lead to diseases such as cancer, immune system defects, osteoporosis and rheumatoid arthritis (Tastan Bishop and Kroon, 2011). In parasites they have several diverse functions. They allow parasites to bore through cellular and tissue barriers, degrade host proteins, manipulate the host immune system and elude host immune responses (Atkinson et al., 2009; Sajid and McKerrow, 2002; Tastan Bishop and Kroon, 2011). Inhibition of *P. falciparum* cysteine peptidases has a marked affect on reducing *P. falciparum* merozoite invasion of erythrocytes (Olaya and Wasserman, 1991), demonstrating the role cysteine peptidase play in cell invasion. Furthermore, *T. cruzi's* major cysteine peptidase *Tcr*CATL, has been linked to plasma leakage in post-capillary venules and may recruit macrophages for invasion (Svensjo et al., 1997).

For cysteine peptidases there are ten clans recognised (CA, CD, CE, CF, CH, CL, CM, CN, CO and C- family of cysteine peptidases not assigned to a particular clan (Caffrey and Steverding, 2009). Clan CA alone contains 20 families (Barrett and Rawlings, 2004). The C1 family, which will be the focus of the present study, contains the archetypal plant cysteine peptidase, papain, and the mammalian cathepsins B, C, K, L and S (Caffrey and Steverding, 2009).

Clan CA cysteine peptidases possess highly conserved cysteine, histidine and asparagine residues making up the catalytic triad found in the active site. These residues are embedded in highly conserved peptide motifs, CGS**C**WAFS, **H**(G/A)VXXVG(Y/W) and YWXIX**N**SWXXWG (triad residues in bold) (Lecaille et al., 2002). Other conserved residues include the ERF/WNIN motif which is characteristic of the CATL-like subfamily (Karrer et al., 1993). The motif which is absent in CATB-like peptidases, is, however, found in *Trypanosoma* cysteine peptidases, with the last two residues in the motif replaced by alanine (Caffrey et al., 2001). The catalytic site of these peptidases is located within the substrate binding region which has a set of subsites or binding pockets (Lecaille et al., 2002; Smooker et al., 2010). These subsites bind the substrate amino acid residues in the N- and C-terminal direction of the scissile bond. The subsites in the N-terminal direction are termed S_1, S_2, S_3…Sn and those in the C-terminal direction are termed S_1', S_2', S_3'…Sn' (Schechter and Berger, 1967) (Fig. 1.5). The corresponding substrate (or inhibitor) amino acids residues binding to the specific subsites are termed P_1, P_2, P_3…Pn and P_1', P_2', P_3'…Pn'. Members of the clan CA have a substrate specificity defined by the S2 pocket, with hydrolysis of small peptides that contain Phe in P_2 typical of CATL and

those with Arg in P_2 typical of CATB (Barrett and Rawlings, 2004; Sajid and McKerrow, 2002).

CATL, CATB (Caffrey and Steverding, 2009) and papain-like peptidase (Lecaille et al., 2002) translation products comprise of a signal peptide, a propeptide and a catalytic domain with the latter representing the mature proteolytically active enzyme. The propeptide is multifunctional. It acts as a scaffold for protein folding of the catalytic domain (Schilling et al., 2001), operates as a chaperone for transport of the proenzyme to the endolysosomal compartment (Nissler et al., 1998) and acts as a high-affinity reversible inhibitor preventing early activation of the catalytic domain (Billington et al., 2000; Guo et al., 2000).

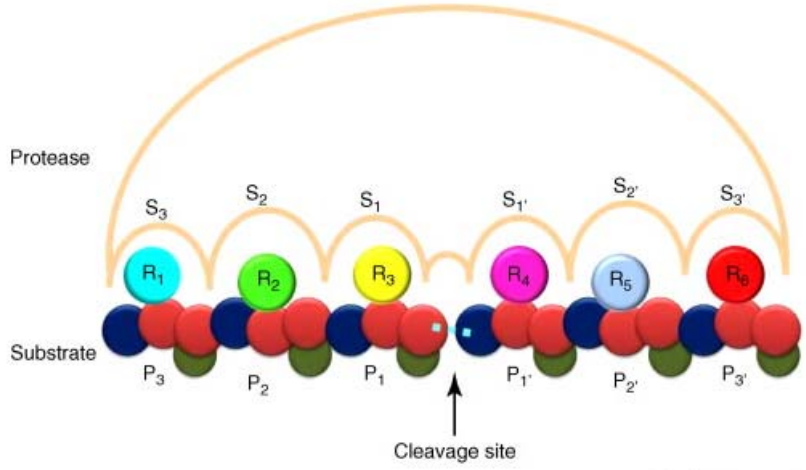


Figure 1.5 Schematic diagram of a peptidase active site based on the Schechter and Berger notation (Smooker et al., 2010).

The overall 3-dimensional fold of cathepsin and papain-like cysteine peptidases is highly conserved. The structure consists of an L and R domain with a central helix and a β-barrel motif being the most prominent features respectively (Fig. 1.6) (Turk et al., 2000). Two of the three residues forming the catalytic triad, Cys and His, are positioned at the N-terminus of the central helix of the L domain and the part of the β-barrel structure of the R-domain respectively (Turk et al., 2000). The two-domain interface forms a ‘V’-shaped active site cleft.

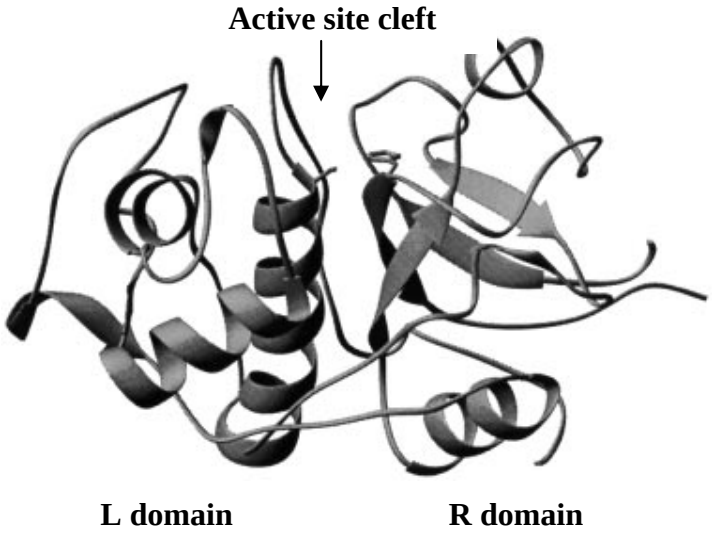


Figure 1.6. Ribbon fold of mammalian CATL. CATL-like cysteine peptidases follow a common papain 3-dimensional framework with an L and R domain and V-shaped active site cleft. The active site cleft is where the catalytic triad interacts with the substrate (Guncar et al., 1999).

1.6.2 General mechanism of hydrolysis by cysteine peptidases

In the cysteine peptidase the nucleophile that attacks the scissile peptide bond is the sulfhydryl group of a cysteine residue (Barrett and Rawlings, 2004). The thiol group is enhanced as a nucleophile due to the close proximity of an active site histidine which acts as a proton donor/general base (Sajid and McKerrow, 2002) (Fig. 1.7). The sulfhydryl group and imidazole of cysteine and histidine respectively give rise to the thiolate-imidazolium charged diad. This is often stabilised by the highly conserved asparagine (Sajid and McKerrow, 2002). During the hydrolysis of a peptide the sulfhydryl cysteine attacks the carbonyl carbon of the scissile bond of the bound substrate and forms a tetrahedral intermediate (Fig 1.7) (Lecaille et al., 2002) which is generally stabilised by the oxyanion hole. The highly conserved glutamine comprises the oxyanion hole and is very important as an electrophilic centre to stabilise the tetrahedral intermediate (Sajid and McKerrow, 2002). Following acylation, release of the N-terminal portion of the substrate, restoration of the deprotonated histidine and formation of thioester bond, deacylation occurs, hydrolysing the thioester bond and releasing carboxylic acid and free enzyme (Fig. 1.7).

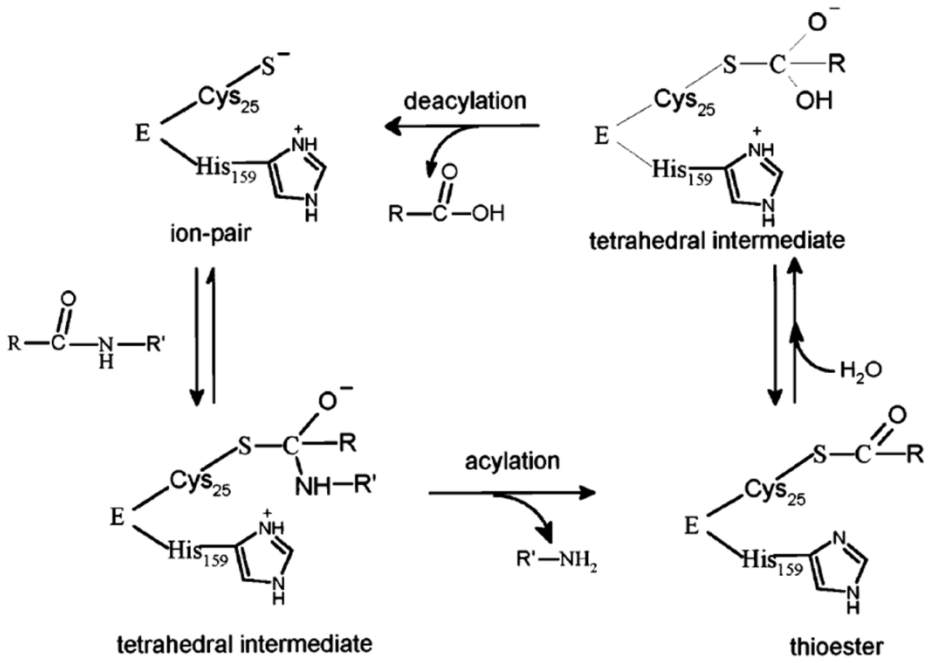


Figure 1.7 Reaction mechanism for cysteine peptidase by means of the catalytic triad (Lecaille et al., 2002)

1.6.3 Parasite cysteine peptidases

With regards to African trypanosomes, three major CATL-like cysteine peptidases have been studied, *Tbr*CATL, *Tbb*CATL and *Tco*CATL. *Tvi*CATL, the major CATL-like cysteine peptidase from *T. vivax* has been studied to a much lesser extent. *Tcr*CATL, from *T. cruzi*, which is the most advanced in pre-clinical investigations for target-based drug development in curing Chaga's disease, shares 59% and 45% identity with *Tbb*CATL and mammalian CATL respectively (Eakin et al., 1992) and 57% with *Tco*CATL. Like mammalian CATL, the major trypanosomal cysteine peptidases are synthesised as pre-pro-proteins (Cazzulo et al., 1989; Eakin et al., 1992; Sakanari et al., 1997).

However, unlike their mammalian homologs, an additional 11-13 kDa C-terminus extension, of unknown function, is separated from the mature catalytic domain of trypanosomal cysteine peptidases by a poly-threonine (*Tcr*CATL and *Tvi*CATL) or poly-proline (*Tbb*CATL and *Tco*CATL) sequence (Caffrey et al., 2000). Parasite cysteine peptidases also have a Cys-His-Asn catalytic triad (Caffrey and Steverding, 2009; Sajid and McKerrow, 2002).

1.6.3.1 Physiological functions of parasite cysteine peptidases

Parasite cysteine peptidases are virulence factors and have been shown to contribute to disease pathogenesis (Caffrey and Steverding, 2009). RNAi against *Tbb*CATL reduced crossing of the blood brain barrier in an *in vitro* model (Abdulla et al., 2008). In the case of the widely studied *T. cruzi*, amastigote replication, intracellular development and host cell invasion have been implicated with *Tcr*CATL function (Aparicio et al., 2004). Furthermore, the CATL-like peptidases from related parasites, *Leishmania* spp., have also been shown to be involved in cellular invasion. Knockout of nineteen genes of the *lmcpb* tandem array from *L. mexicana*, coding for nineteen highly related CATL-like peptidases, resulted in a decrease of parasite virulence (Frame et al., 2000) with decline in percentage of infected host cells (Mottram et al., 1996). Metacaspases, part of clan CD, which contain a histidine and cysteine catalytic diad, are important virulence factors too. *T. b. brucei* possess three metacaspase genes. Helms et al (2006) demonstrated, through gene deletion, that three of the of the metacaspases, MCA2, MCA3 and MCA5 are involved in recycling processes of RAB11-positive endosomes. This recycling is, however, independent of VSG recycling.

*Tco*CATL (Mbawa et al., 1992) and *Tbb*CATL (Troeberg et al., 1996) have been shown to degrade a range of proteins, such as bovine fibrinogen, bovine serum albumin and ovalbumin at physiological pH *in vitro*. The cysteine peptidase from *Trichomonas vaginalis*, the causative agent of trichomoniasis, displayed activity against collagen IV, fibronectin and haemoglobin (Mendoza-Lopez et al., 2000) Additionally, falcipain-2, a CATL-like cysteine peptidase from *Plasmodium falciparum*, cleaves erythrocyte skeleton proteins such as ankyrin (Dua et al., 2001). *T. b. brucei* type I PGP (pyroglutamyl peptidase), belonging to the C15 family of cysteine peptidases, removed the N-terminal pyroglutamic acid blocking groups from thyrotrophin-releasing hormone and gonadotropin-releasing hormone, dramatically reducing the half-life of these neural peptides (Morty et al., 2006).

*Tbb*CATL (Troeberg et al., 1996), *Tcr*CATL (Chagas et al., 1997) and *Tco*CATL are active a neutral pH allowing the peptidases to interact with host proteins in a physiological environment (Lalmanach et al., 2002). Furthermore, the recombinant catalytic domain of *Tvi*CATL is also active close to physiological pH (Vather, 2010). Although there are

natural endogenous host cysteine peptidase inhibitors, incomplete enzyme inhibition with such inhibitors has been documented (Lonsdale-Eccles et al., 1995; Wiser et al., 1997). In some instances, biologically active molecules have been released from such inhibition and as a result may play a role in pathology. *Tcr*CATL has been able to display kininogenase activity *in vitro* (Del Nery et al., 1997). Initial work on *Tco*CATL interaction with bovine kininogens indicated that the interaction does not achieve complete enzyme inhibition and triggers the release of kinins (Lalmanach et al., 2002).

1.6.3.2 Biochemical characterisation of parasite CATL peptidases

Clan CA peptidases are characterised by their sensitivity to the general cysteine peptidase inhibitor E-64 and by having substrate specificity defined by the S_2 pocket (Sajid and McKerrow, 2002). *Tco*CATL has a distinct preference for the presence of a basic amino acid (Arg of Lys) in the P_1 position (Mbawa et al., 1992). Furthermore, *Tco*CATL hydrolyses peptides with hydrophobic or aromatic side chains, such as Phe or Leu, in the P_2 position. These are features common to many protozoan cysteine peptidases such as the type I CATL-like peptidases from *Leishmania* (Robertson and Coombs, 1990), *Tcr*CATL (Chagas et al., 1997), *Tbb*CATL (Troeberg et al., 1996), *Tbr*CATL (Caffrey et al., 2001) and falcipain-2 (Shenai et al., 2000). These CATL-like peptidases, besides *Tcr*CATL, poorly cleave the characteristic CATB substrate Z-Arg-Arg-AMC. The ability of CATB to hydrolyse Abz-Arg-Arg-AMC is due to the Glu residue at position 205 (papain numbering), also found in *Tcr*CATL, lining the S_2 subsite. The inability of *Tco*CATL to hydrolyse the CATB substrate is due to the Leu205 is the S_2 pocket. The substitution of a glutamic acid by a leucyl residue does not allow for the accommodation of Arg in P_2 of *Tco*CATL (Lecaille et al., 2001) and therefore the superimposition model of *Tco*CATL and *Tcr*CATL does argue strongly for a more restricted specificity for *Tco*CATL (Fig. 1.8). Mutating the Gly, lining the S_2 pocket, to Glu in *Lma*CATB, a CATB peptidase possessing CATL-like substrate specificity, allowed for cleavage of Z-Arg-Arg-AMC (Chan et al., 1999). Both *Tco*CATL (Mbawa et al., 1992) and *Tbr*CATL (Troeberg et al., 1996) have demonstrated the ability to cleave CATL-like substrates with a Pro, Val or Ala in the P_3 position

1.6.3.3 Trypanosomal peptidases as targets for drug design

To date, *Tcr*CATL (McGrath et al., 1995) and *Tbr*CATL (Kerr et al., 2009) are the only trypanosomal CATL-like cysteine peptidases that have been crystallised and their structures solved. This advantage has allowed for the directed design of synthetic peptidase inhibitors. Inhibitors, such as the irreversible dipeptidyl substrate-analogues coupled to the reactive fluoromethylketone (CH_2F, FMK) or diazomethylketone (CHN_2, CMK) groups have been widely researched for selectively inhibiting the parasitic cysteine peptidases. A number of chloromethylketones, fluoromethylketones and diazomethylketones (DMK), which inhibit cysteine peptidases, are able to kill bloodstream forms of *T. b. brucei in vitro* (Troeberg et al., 1999).

Furthermore, Mbawa et al. (1992) reported that some peptidyldiazomethylketones (Z-Leu-Leu-Met-CHN_2, Z-Leu-Met-CHN_2 and Z-Leu-Lys-CHN_2) were trypanocidal when tested against *T. congolense in vitro*. However a new generation of inhibitors, the vinyl sulfone group (Palmer et al., 1995), that have a peptidomimetic, rather than a peptidyl backbone, have better selectivity for cysteine peptidases over serine peptidases and stable inactivation of the target enzyme (Kerr et al., 2009).

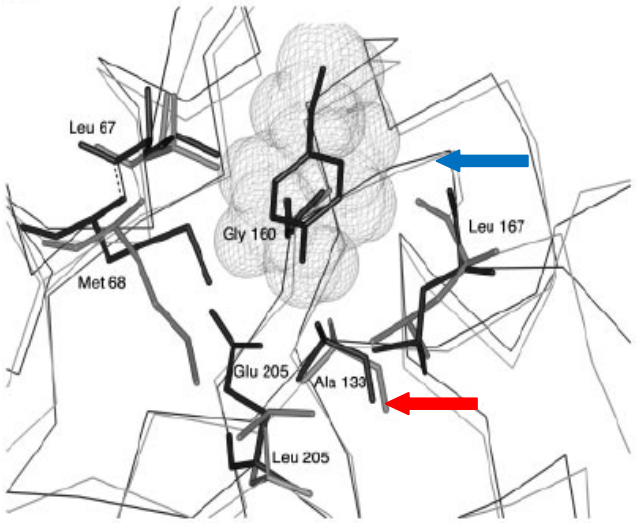


Figure 1.8 Structure of the S_2 subsite of *Tco*CATL and *Tcr*CATL. Structure of the S2 subsite of *Tco*CATL (grey). The S2 subsite of *Tcr*CATL (black), corresponding to the cruzipain/Z-Tyr-Ala-FMK complex, is superimposed. The red arrow indicates the residue difference, Leu205 for congopain and Glu 205 for *Tcr*CATL. For the sake of clarity, only the tyrosyl residue of Z-Tyr-Ala-FMK is represented (blue arrow). Residue numbers correspond to papain numbering (Lecaille et al., 2001).

The peptidomimetic vinyl sulfone, K11777, which targets *Tcr*CATL, is lethal to *T. cruzi* parasites both *in vitro* and in animal models of disease and is an example of 'chemical knockout' of a the *Tcr*CATL gene (Engel et al., 1998a; Engel et al., 1998b). Furthermore, the vinyl sulfone K11777 was very effective in killing bloodstream forms of *T. b. brucei* (Troeberg et al., 1999). The inhibitor, which has no significant toxicity to rodent or canine hosts (Abdulla et al., 2007), is currently completing pre-clinical testing for the treatment of Chaga's disease. It is imperative that the crystal structures of pathogenic factors be determined to provide information to help guide the design of specific drugs.

1.7 Protein crystallisation

In order for protein crystals to form, ordered clusters of protein molecules, nuclei, need to form. The level of supersaturation, as shown by the supersaturated regions of the phase diagram (Fig. 1.9), must be high enough to initiate nucleation and then support crystal growth (McPherson et al., 1995). However higher levels of supersaturation are required for nuclei to form than for subsequent crystal growth. The level of supersaturation for spontaneous nucleation to occur and for crystal growth support is known as the *labile* and *metastable zone* respectively (Luft and DeTitta, 2009) (Fig. 1.9). As protein crystals grow, more protein is sequestered from the surrounding solution, causing the concentration of protein in the surrounding solution to decrease. As a result of crystal growth, the level of supersaturation is continually lowered until a state of equilibrium between the crystal (a solid phase) and the surrounding solution (liquid phase) is reached (Luft and DeTitta, 2009). The solid phase, the protein crystal, will redissolve if the solution becomes undersaturated.

The sequence of steps involved in the crystallisation process is shown in Fig. 1.10. Preliminary protein preparation, prior information and characterisation are critical to crystallisation trials. Recombinant DNA technology provides a means of making specific changes to protein character to improve the chances of crystallisation (Gilliland and Ladner, 1996). Tags in recombinant protein constructs has allowed for efficiently yielding

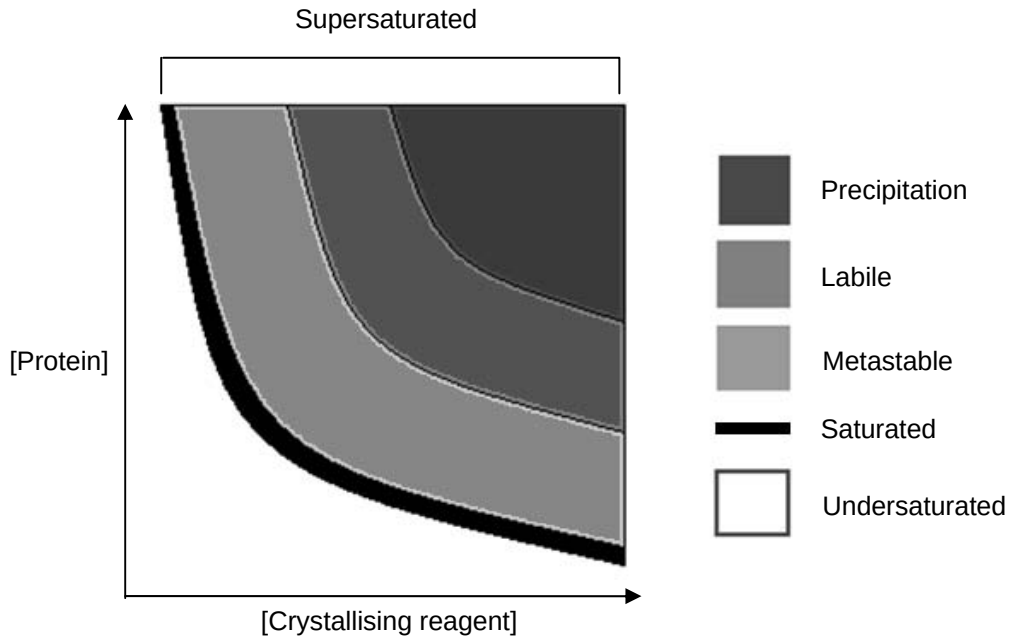


Figure 1.9 Phase diagram for protein crystallisation. The y axis is the macromolecule concentration and the x axis is the crystallising reagent concentration. The crystallising reagent is a chemical whose concentration is inversely related to the macromolecule solubility. The supersaturated region is divided in three zones, the precipitation, labile and metastable. The thick dark line represents the limit of solubility, where below it the solution is undersaturated (Luft and DeTitta, 2009).

a large amount of highly purified protein for crystallisation (Rupp, 2009). Furthermore, altering protein sequence and mutating specific residues in a protein sequence have improved recombinant protein solubility and stability in several instances. The protein concentration, quoted in the Protein Data Bank, for successful crystallisation is 14 mg/ml. However, proteins have been crystallised at much lower concentrations and as Rupp (2009) describes as a general guideline, any concentrations from 2-10 mg/ml is sufficient.

The choice of crystallisation technique also plays a role in the success of crystallisation. It has been demonstrated that drop shape and volume as well as distance of drop from the reservoir affects the results (Luft et al., 1996). The DST-NRF financially supported Rigaku crystallisation robotic suite, at the University of KwaZulu-Natal, uses the sitting drop vapour diffusion technique, where the vapour-phase equilibration between an undersaturated protein solution and a dehydrating solution leads to a supersaturated state. Other techniques include batch and liquid diffusion. The approach to supersaturation

changes with the crystallisation method as each method has a unique path through the phase diagram (Luft and DeTitta, 2009).

Random screening or custom screening can be used for initial screens. Protein solubility screens can be performed prior to crystallisation screens in order to determine the appropriate protein concentration for crystallisation. This generally involves probing the solubility of proteins at different concentrations in different buffers and crystallising reagents to produce a supersaturation curve of the protein (Benvenuti and Mangani, 2007). In this way, screens can be designed to reach the metastable zone. There are also commercially available kits to test the solubility of proteins. If limited screening is chosen initially, conditions need to be carefully decided on in order to maximise the chances of getting a crystallisation hit. If there are crystallisation conditions available of related previously crystallised proteins (similar enzyme character and sequence), then a screen design may be based on these conditions. However, there is the other approach of fast screening using commercial screens that have a large screening set and are said to reduce or remove the cycles of optimisation later (Newman, 2009). If screening, even after optimisation, produces no positive results, protein constructs need to be reconsidered. Diffracting protein crystals can be used for structure determination. However, non-diffracting crystals require further optimisation of crystallisation conditions.

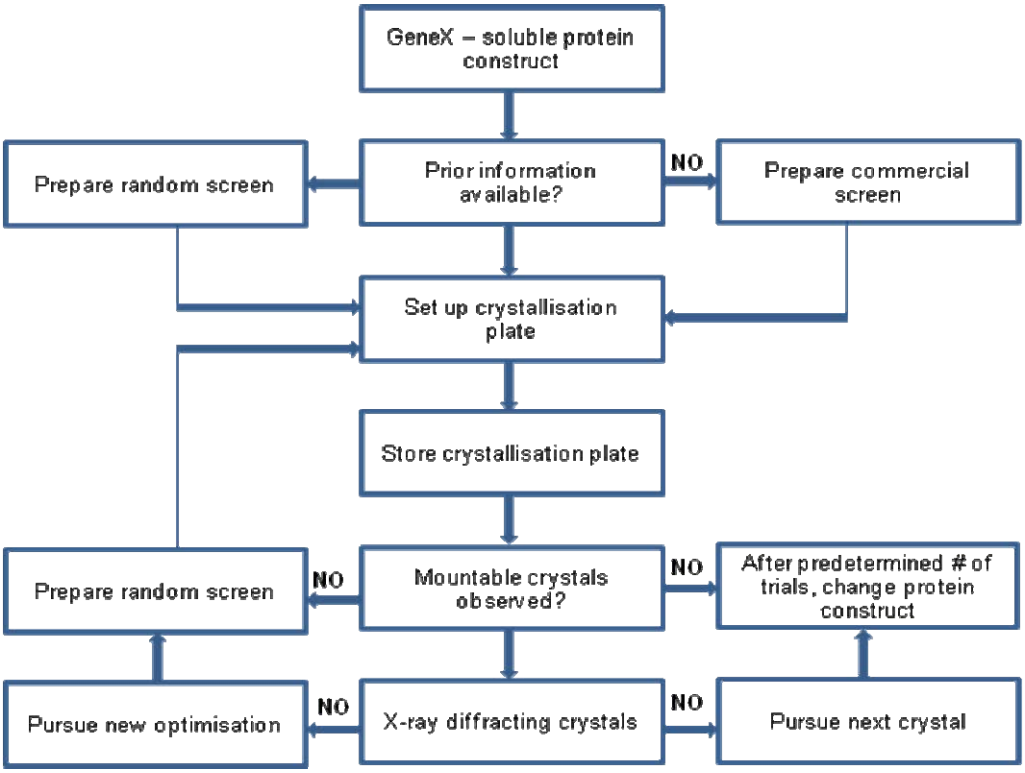


Figure 1.10 Sequence of steps in a crystallisation project. Adapted from Rupp and Wang (2004).

1.7.1 Parasitic cysteine peptidase crystal structures

As mentioned earlier, the overall 3-dimensional fold of cathepsin and papain-like cysteine peptidases is highly conserved consisting of an L and R domain with a central helix and a β-barrel motif respectively. Charged buried residues that form electrostatic interactions at the interface of the two domains in papain, and which are conserved in all known sequences of trypanosomal cysteine peptidases, are involved in the catalytic mechanism of cysteine peptidases. These interdomain electrostatic interactions, which are also found in mammalian cathepsins H and L, have been conserved during evolution to ensure proper folding and stability of cysteine peptidases of the papain family (Kamphuis et al., 1985).

*Tcr*CATL was the first trypanosomal cysteine peptidase to be crystallised. The refined crystal structure of *Tcr*CATL in complex with Z-Phe-Ala-FMK was compared to the structure of papain (McGrath et al., 1995). The main fold comprises of an L and R domain with an active-site cleft. Certain helices in the L domain differ from papain, however, the antiparallel β-structure of the R domain as well as the catalytic site is isostructural for papain and *Tcr*CATL (McGrath et al., 1995). An extra helix is created from residues of the C-terminus of *Tcr*CATL. Superimposition of *Tcr*CATL (in complex with Z-Phe-Ala-

FMK) with inhibitor-bound papain and free papain showed that *Tcr*CATL is similar in structure to inhibitor-bound papain (McGrath et al., 1995). Therefore, the inter-domain cleft widens upon binding of inhibitor.

The predicted 3D structure of *Tco*CATL, based on the sequence identity with *Tcr*CATL, follows a common framework of an L and R domain with an active-site cleft found for papain-like cysteine peptidases (Fig 1.11). The catalytic triad, Cys 25, His 159 and Asn 175, and the residues defining the extended substrate-binding site are located in the cleft between the R domain and the L domain (Lecaille et al., 2001). Although there is a homology based model of *Tco*CATL, *Tco*CATL has been demonstrated to dimerise. The dimeric formation is yet to be determined (Ndlovu, 2010).

The common papain two-domain fold is also evident for *Tbr*CATL (Kerr et al., 2009). The commonality between *Tcr*CATL and *Tbr*CATL in the papain-like fold allowed for superimposing the two structures. There is a striking difference in the S_1' subsite of the two enzymes. The phenyl-sulfone group in the P_1' position of the K11777 inhibitor in complex with *Tbr*CATL is flipped approximately 90° out of the active site in relation to the *Tcr*CATL-K11777 complex (Fig. 1.11) (Kerr et al., 2009). The substitution of Trp for the slightly less bulky Phe at position 144 in *Tbr*CATL allows the amino acid residue to more readily access a deep, buried pocket in the bottom of the S_1' subsite, consequently allowing the neighbouring Met 145 to penetrate deeper into the S_1' subsite (Kerr et al., 2009). This prevents the phenyl sulfone of K11777 to lie flat in the subsite.

The polypeptide chain of the mature CATL-like cysteine peptidase from *P. falciparum*, falcipain-2, folds as a prototypical structure of papain-like proteases, with two distinct domains separated by a long central substrate binding cleft containing the active site (Hogg et al., 2006). The L domain is predominantly α-helical (α2, α3, α4) with several long disulfide-stabilised segments lacking regular secondary structure, whereas the R domain contains a large antiparallel β-sheet (β2-β7-β3-β4-β5-β6). Additional structural features that are unique to the falcipain-2 subfamily (with members such as the vivapain, from *P. vivax* and berghepain, *P. bergheii* homologues) include a 17-residue N-terminal extension and a 14-residue β-hairpin protuberance extending from β4 and β5 that comprises the hemoglobin binding motif (Hogg et al., 2006). A comparative structural analysis of the

falcipain-2 active site with that of *Tcr*CATL suggests that existing vinyl sulfone inhibitors designed against cysteine peptidase may be useful starting compounds for rational design of novel antimalarials.

The CATL-like cysteine peptidases from the parasitic liver fluke, *Fasciola hepatica*, FheproCL1Gly^{25}, with the replacement of Cys 25 with Gly, also comprises of a mature domain which is bi-lobed and substrate-binding cleft running between the two lobes of the enzyme, characteristic of the papain fold (Stack et al., 2008). However, FheproCL1Gly^{25} has unique "gatekeeper" residues 157 and 158 at the entrance of the active-site pocket. Similarly, falcipain-3 from *P. falciparum*, is restricted for a CATL-like peptidase through a combination of two "gatekeeper" residues at the bottom of the S_2 subsite (Kerr et al., 2009). Falcipain-3 does share the common framework of papain-like cysteine peptidases. However, as was reported that falcipain-3 deviates slightly from the classical papain fold in that there are two insertions on either terminus, unique to plasmodial cysteine peptidases (Kerr et al., 2009).

1.8 Objectives of present study

Trypanosomal cysteine peptidases are of major pathological importance. It is therefore vital that such peptidases be structurally and functionally characterised. CATL-like cysteine peptidases from *T. congolense* (Mbawa et al., 1992), *T. cruzi* (Cazzulo et al., 1990), *T. b. rhodesiense* (Caffrey et al., 2001) and *T. b. brucei* (Troeberg et al., 1996) have been previously characterised. However, no cysteine peptidase has to date been characterised from *T. vivax*. CATL-like sequences from African and South American isolates of *T. vivax* have been evaluated for their suitability as genetic markers for population structure analysis and diagnosis (Cortez et al., 2009).

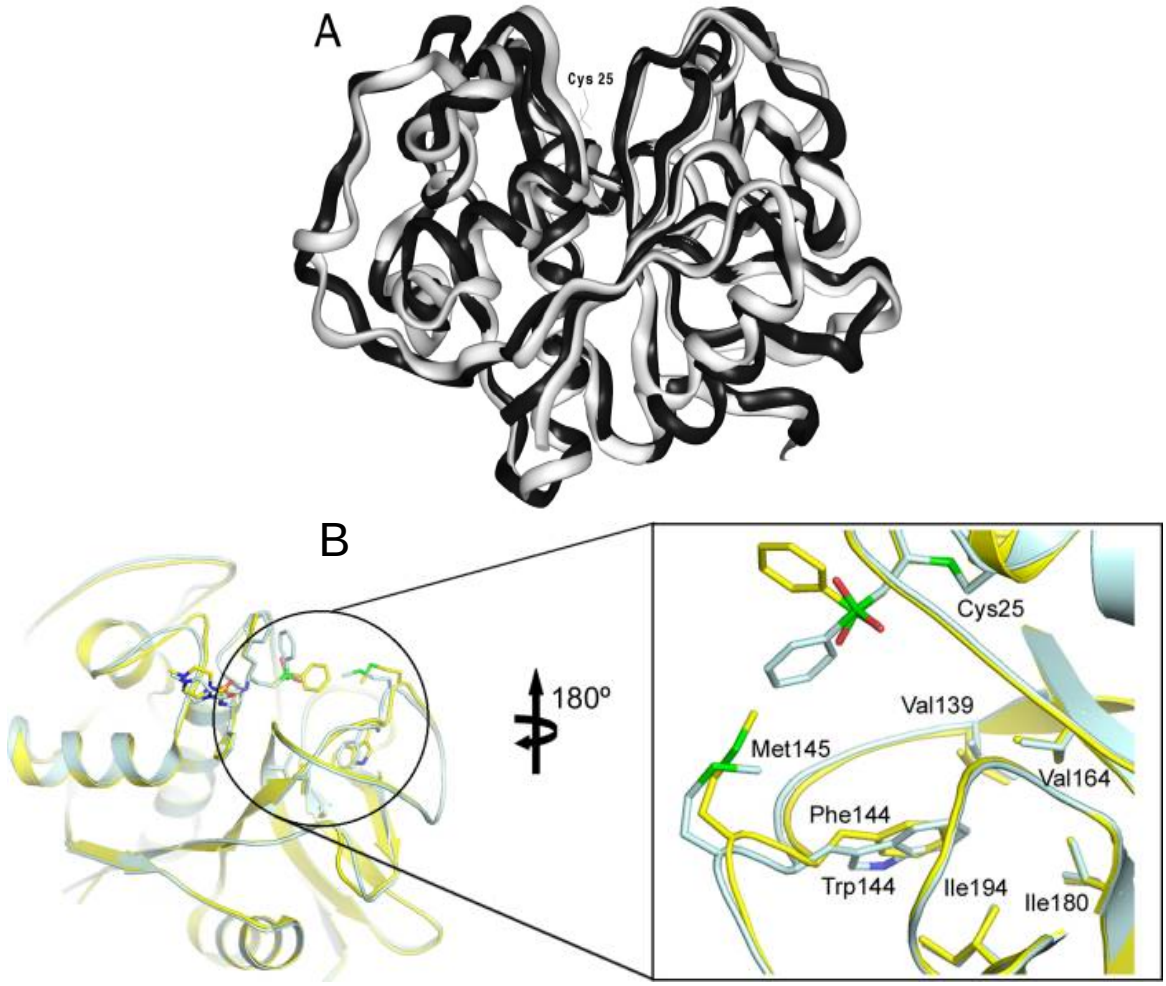


Figure 1.11 The ribbon superimposition model of *Tco*CATL and phenyl sulfone flipping in *Tbr*CATL. (A) The ribbon plot superimposition of a homology-based *Tco*CATL (grey) model (catalytic domain) and of *Tcr*CATL (black). The Cys25 residue of the catalytic triad is indicated in the cleft between the R and L domain (Lalmanach et al., 2002). (B) Ribbon representation and detailed inlay of the S1' subsite of *Tcr*CATL (pale blue) and *Tbr*CATL (yellow). Important residues in subsite opening are depicted in ball and stick representation. Bound ligands are coloured as their respective peptidase partners (Kerr et al., 2009).

The analysis of the CATL-like amino acid sequence in the present study did reveal that there were distinct regions in the *T. vivax* CATL-like sequences, which like other cysteine peptidases could be inferred to represent the pre-, pro-, central catalytic domain (with the cysteine, histidine and asparagine catalytic triad) and C-terminal extension. Determining whether *T. vivax* CATL-like peptidases are pathogenic factors and similar in structure, function and character to other pathogenic trypanosomal CATL-like peptidases would be pivotal to the design of a general inhibitor for trypanosomal cysteine peptidases.

In the present study, native *Tvi*CATL was purified from bloodstream *T. vivax* strain 486 and biochemically characterised for substrate specificity, pH optimum, ability to digest bovine fibrinogen and inhibition studies using both reversible are irreversible inhibitors. The results obtained for this part of the study were presented in Chapter 2.

Trypanotolerant cattle infected with *T. congolense* have shown to develop prominent IgG responses to *Tco*CATL (Authié et al., 1993). It is for this reason that *Tco*CATL is being evaluated as an anti-disease vaccine candidate to reduce the symptoms of the disease. Understanding the dimeric conformation of *Tco*CATL will facilitate improving antigenicity and mapping *Tco*CATL epitopes, important for vaccine design. Furthermore, screening inhibitors that are structurally suited for *Tco*CATL is important for inhibitor design. To date, *Tcr*CATL (McGrath et al., 1995) and *Tbr*CATL (Kerr et al., 2009) are the only trypanosomal CATL-like cysteine peptidases that have been crystallised and structures solved. This advantage has allowed for the directed design of synthetic peptidase inhibitors.

In the second half of the study the catalytic domains of *Tco*CATL and *Tvi*CATL, as well as the full-length PGP, were recombinantly expressed and purified. These proteins along with *Tco*CATL dimerisation mutants, *Tco*CATL (H43W), *Tco*CATL (K39F; E44P) and the proenzyme mutant *Tco*CATL (C25A) were screened for a suitable crystallisation condition using the Rigaku robotic crystallisation suite. Crystallisation hits for *Tco*CATL (K39F; E44P) were further optimised. The results of these studies are presented in Chapter 3. Finally the results obtained in the present study are summarised and discussed in Chapter 4.

CHAPTER 2
PURIFICATION AND CHARACTERISATION OF NATIVE *Tvi*CATL FROM *Trypanosoma vivax*

2.1 INTRODUCTION

Livestock trypanosomosis (Nagana) is a chronic disease with major economic impact in tropical countries. The disease is caused by tsetse-transmitted haemoprotozoa *Trypanosoma congolense*, *T. vivax* and *T. brucei brucei*. Nagana causes estimated losses to African agriculture of US $4.75 billion per year (Van den Bossche et al., 2010), affecting about 10 million km^2 of sub-Saharan Africa (Kabayo, 2002). Currently there is no vaccine for trypanosomosis and although a pan-African tsetse fly and trypanosomosis eradication programme, under the auspices of the Organisation of African Unity (OAU), has been launched (Lalmanach et al., 2002), studies aimed at identifying potential targets for new trypanocidal drugs have been running concurrently. The three compounds, diaminazene, isometamidium and ethidium have been used for more than 50 years (Delespaux et al., 2010) and as a consequence drug resistance of *T. congolense* has been reported in 17 countries of sub-Saharan Africa (Delespaux et al., 2008). These compounds are thus no longer effective for large and small-scale treatment.

There are many hurdles facing the development of a vaccine against trypanosomosis. One of the major obstacles is the phenomenon of antigenic variation that trypanosomes possess. Trypanosomes are able to switch their variant surface glycoprotein (VSG) coat, giving rise to a new antigenically distinct trypanosome population (Donelson, 2003), thereby evading the host immune system. *Tco*CATL, the major CATL-like cysteine peptidases from *T. congolense*, has been shown to act as a pathogenic factor in *T. congolense*-infected cattle (Authié et al., 2001). In addition, trypanotolerant cattle infected with *T. congolense* were found to develop prominent IgG responses to *Tco*CATL (Authié et al., 1993). Consequently *Tco*CATL is being evaluated as an “anti-disease” vaccine candidate with the aim to reduce the symptoms of the disease, rather than eliminating the parasite.

It has previously been established that parasite peptidases are promising chemotherapeutic or vaccine targets (McKerrow, 1999; McKerrow et al., 1999). Besides the many roles that such peptidases play, further attention has focused on the clan CA peptidases with family C1 (CATB and CATL-like) peptidases being of major interest. Parasitic CATL-like cysteine peptidases are critical to the life-cycle and pathogenocity of many parasites (Sajid and McKerrow, 2002). In the case of the widely studied *T. cruzi*, amastigote replication, intracellular development (Franke de Cazzulo et al., 1994) and host cell invasion (Aparicio et al., 2004) have been implicated with *Tcr*CATL function. *Tco*CATL (Mbawa et al., 1992) and *Tbb*CATL (Troeberg et al., 1996) degrade a range of proteins *in vitro*.

Like *Tbb*CATL and *Tcr*CATL, *Tco*CATL is active at neutral pH allowing it to interact with host proteins in a physiological environment (Lalmanach et al., 2002). Furthermore, the recombinant catalytic domain of *Tvi*CATL is also active close to physiological pH (Vather, 2010). General substrate specificity for these CATL-like peptidases is defined by the S_2 pocket (Fig. 2.1). They are capable of hydrolysing peptides with a hydrophobic or aromatic residue in the P_2 position. Although there are physiological natural cysteine peptidase inhibitors, incomplete enzyme inhibition with such inhibitors has been documented (Lonsdale-Eccles et al., 1995; Wiser et al., 1997). In some instances, biologically active molecules have been released from such inhibition and as a result may play a role in pathology. *Tcr*CATL has been able to display kininogenase activity *in vitro* (Del Nery et al., 1997). The structure/function relationship and catalytic mechanism of this C1 peptidase family is known, which provides a foundation for the synthetic and high throughput computational drug design based of ligand screening (Sajid and McKerrow, 2002). Adding to this, K11777, a selective vinyl sulfone-derivitised pseudopeptide inhibitor of *Tcr*CATL, which has undergone rodent and dog toxicology tests, (McKerrow et al., 2008), was effective in curing or alleviating *T. cruzi* infection in preclinical proof-of-concept studies and has now entered formal preclinical drug development investigation (McKerrow et al., 2009).

CATL-like cysteine peptidases contain a signal peptide, a pro-peptide and a catalytic domain (Sajid and McKerrow, 2002). The catalytic domain consists of the highly

conserved catalytic triad (Cys, His and Asn) (Fig. 2.1). However, 16% of the parasitic protozoa C1 peptidases do not appear to have the classic catalytic triad (Atkinson et al., 2009; Pillay et al., 2010). Kinetoplastid CATL-like peptidases also have a unique highly conserved 11-13 kDa C-terminal extension (Boulangé et al., 2001; Caffrey and Steverding, 2009). The mature catalytic domain is separated from the C-terminal extension by a poly-threonine (*Tcr*CATL and *Tvi*CATL) or poly-proline (*Tbb*CATL and *Tco*CATL) tract. The definitive biological function of the C-terminal extension is not yet known (Caffrey et al., 2000).

Trypanosomal cysteine peptidases from *T. congolense* (Mbawa et al., 1992), *T. cruzi* (Cazzulo et al., 1990), and *T. b. brucei* (Troeberg et al., 1996) have been previously characterised. Additionally a novel cysteine peptidases from *T. evansi* have recently been recognised, which is inhibited by the cysteine peptidase inhibitor E-64, active at acidic pH and retains activity in serum samples of rats infected with *T. evansi* (Yadav et al., 2011). However, to date no cysteine peptidase has been characterised from *T. vivax*. Endopeptidase variation among the different life-cycle stages of West- and East-African *T. vivax* has, however, been previously reported (Mbawa et al., 1992). For diagnostic and genetic marking in population structure analysis, CATL-like sequences, from African and South American isolates of *T. vivax*, have been evaluated (Cortez et al., 2009). Analysis of the *T. vivax* CATL-like amino acid sequences (Fig. 2.1) show that there are typical cysteine peptidase sequences representing the pre-, pro, central catalytic domain (with the cysteine, histidine and asparagine catalytic triad) and a C-terminal extension.

Whether *T. vivax* CATL-like peptidases are pathogenic factors is still to be determined. However, pivotal to the design of a general inhibitor is the confirmation of whether or not *T. vivax* CATL-like peptidases act as pathogenic factors and possess similarity in structure, function and character to other trypanosomal CATL-like peptidases. In the present study the purification and enzymatic characterisation of native *Tvi*CATL from *T. vivax* Y486 is reported. The results showed that native *Tvi*CATL shares similar enzymatic characteristics with those of other major trypanosomal cysteine peptidases, *Tco*CATL, *Tcr*CATL and *Tbb*CATL, as well as the recombinant catalytic domain of *Tvi*CATL.

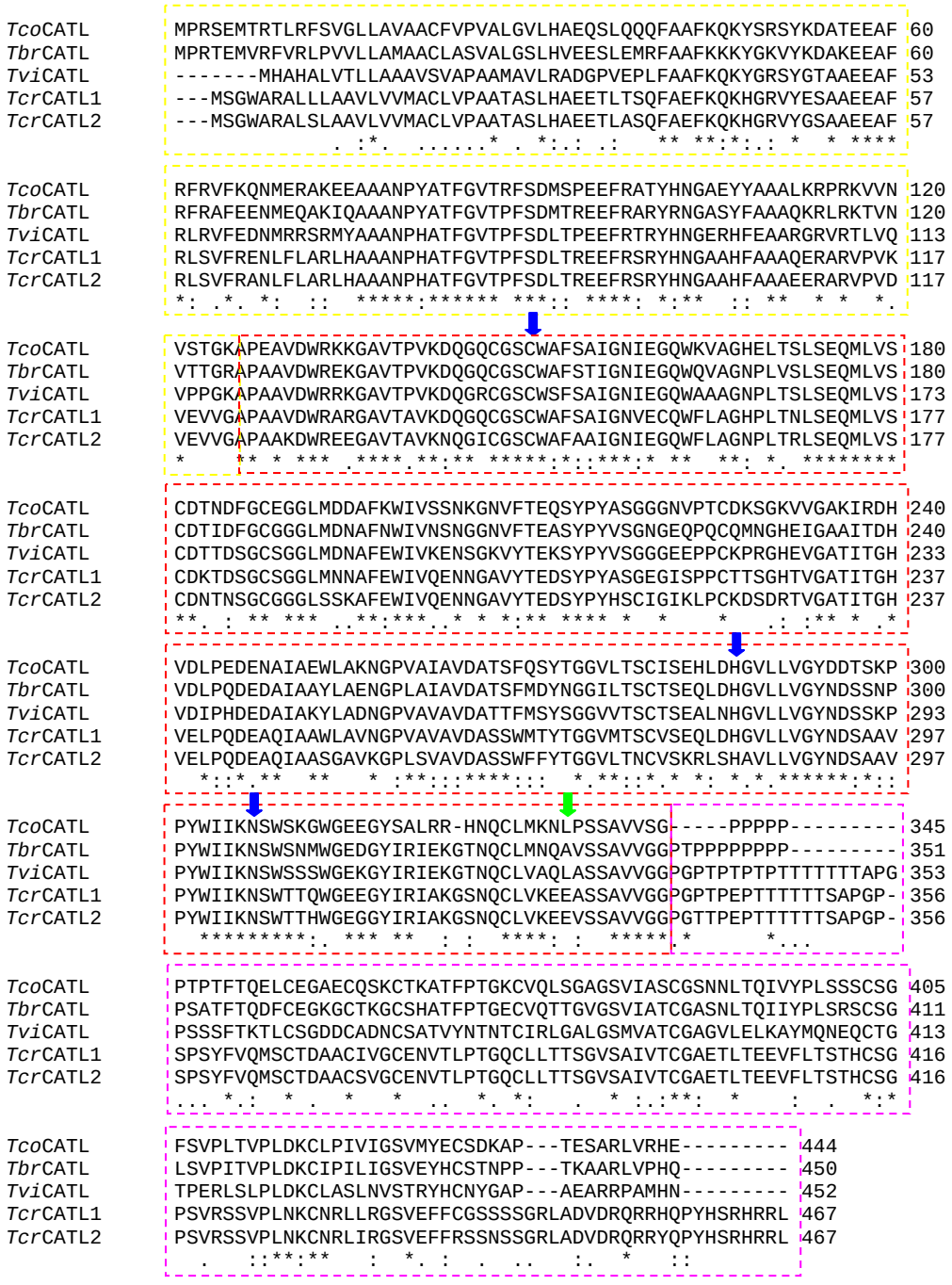


Figure 2.1 Sequence alignment of *Tvi*CATL with other trypanosomal cysteine peptidases. Shown here are the amino acid residues composing part of the signal and propeptide (yellow box), catalytic domain and (red box) C-terminal extension (purple box) of *Tco*CATL (*T. congolense*, EMBL accession no. L25130 (Boulangé et al., 2001), *Tbr*CATL (*T. brucei rhodesiense*, accession no. CAA38238 (Pamer et al., 1990), *Tvi*CATL (*T. vivax*, accession no. CCD21670.1) and *Tcr*CATL 1 and 2 (*T. cruzi*, accession no. M84342 (Eakin et al., 1992) and M90067 (Lima et al., 1994) respectively. Alignment was conducted using ClustalW software (Chenna et al., 2003). Blue arrows indicate the positions of the catalytic triad residues. Green arrow indicates predicted critical residue around the S2 pocket for substrate binding.

2.2 MATERIALS AND METHODS

2.2.1 Materials

Enzyme purification and quantification:

Q-Sepharose, Sephacryl S-200, Cyclohexamide and Histopaque® 1077 were purchased from Sigma (Germany). DE-52 cellulose (DEAE) resin was purchased from Whatman International (England). BCA™ Protein Assay Kit and 10 kDa cut-off Snakeskin® pleated dialysis tubing were from Pierce (Rockford, IL, USA). Nunc-Immuno™ 96-well plates were purchased from Nunc Intermed (Denmark).

Antibodies:

Polyclonal antibodies were previously raised in chickens against the C-terminal extension of *Tvi*CATL and affinity purified (Mkhize, 2003). Rabbit anti-chicken IgY horse raddish peroxidase (HRPO) conjugate was purchased from Jackson Immunochemicals (USA).

Mice:

BALB/c mice were obtained from the Bioresources Unit at the University of KwaZulu-Natal (Westville campus).

Enzyme assays:

The peptide substrates H-D-Val-Leu-Lys-7-amino-4-methylcoumarin (AMC), H-D-Val-Leu-Lys-AMC, H-D-Ala-Leu-Lys-AMC, H-Ala-Phe-Lys-AMC, H-Pro-Phe-Arg-AMC, methoxy-succinyl (MeoSuc)-Asp-Tyr-Met-AMC were obtained from Bachem (King of Prussia, PA, USA). Benzyloxycarbonyl (Z)-Phe-Arg-AMC, Z-Gly-Pro-Arg-AMC, Z-Gly-Gly-Arg-AMC, Z-Arg-AMC, Z-Arg-Arg-AMC, Z-Gly-Arg-Arg-AMC, Z-Ala-Arg-Arg-AMC, E-64, leupeptin, chymostatin, chicken egg white cystatin, antipain, bestatin, ethylenediaminetetra-acetic acid (EDTA), pepstatin A, iodoacetic acid (IAA), iodoacetamide (IAM), phenylmethylsulfonylfluoride (PMSF), N-tosyl-L-lysyl chloromethylketone (TLCK) and N-tosyl-L-phenylalanyl chloromethylketone (TPCK) were obtained from Sigma (Munich, Germany). Nunc Black 96-well plates were purchased from Nunc Intermed (Denmark).

2.2.2 Isolation of *T. vivax*

Cryosuspensions of bloodstream forms of *T. vivax* strain Y486 stabilite were thawed from liquid nitrogen storage and diluted as required with phosphate saline glucose (PSG) [57 mM Na_2HPO_4, 3 mM NaH_2PO_4, 42 mM NaCl, 50 mM glucose and 1mM hypoxanthine]. The viability of the cryosuspended parasites was confirmed by microscopic examination. Two BALB/c mice were injected intraperitoneally with 10^6 parasites. This allowed the development of high parasitaemia in 6 days. The blood parasitaemia was monitored daily by microscopic examination (400x magnification) of blood taken from the tail vein. When high parasitaemia was evident, blood was collected by cardiac puncture and placed into anticoagulant. Approval for procedures involving mice was obtained from the University of KwaZulu-Natal animal ethics committee, reference number 036/09/Animal. Attempts were made to use Wistar rats as an alternate means to expand parasite numbers. However, even after immunosuppressing the rat with 0.5 ml of a 50 mg/ml cyclohexamide, no satisfactory level of parasitaemia was produced following infection with *T. vivax* strain Y486.

Parasites were purified from blood components by anion exchange chromatography on DE-52 resin (Lanham and Godfrey, 1970) followed by Histopaque® 1077 (Sigma, Munich, Germany). Briefly 20 g of DE-52 resin was suspended in 80 ml of phosphate buffered saline (PBS) [100 mM Na_2HPO_4, 2 mM KH_2PO_4, 2.7 mM KCl and 137 mM NaCl, pH 8.0). The resin was allowed to settle in a column (30 x 120 mm) overnight at 4°C. The column was washed with 1 column volume of PBS followed by equilibration with one column of PSG. The infected mouse blood was added to the column with the tap closed to ensure that the blood settled. The tap was then opened and the eluate was collected as 1 ml fractions. Trypanosomes generally pass through the gel unhindered and erythrocytes and other blood components bind to the gel (Lanham and Godfrey, 1970). Parasites were detected in the eluate by placing a drop of eluate on a microscope slide and examining it under a microscope (400 x magnification). PSG was continually added to the column until no more parasites were present in the eluate. To remove any contaminating erythrocytes or blood components, Histopaque® 1077 was used in a 1:1 ratio with the eluant and centrifuged (400 x *g*, 30 min, RT). The opaque trypanosome layer was carefully aspirated, parasites counted using a haemocytometer (Agar Scientific, England) and stored in 70% (v/v) PSG, 30% (v/v) glycerol at -80°C until further use.

2.2.3 Purification of *Tvi*CATL

Lysing parasites:

Parasites (approx. 8.5 x 10^7) were suspended in ice cold 0.1% (w/v) Brij-35 (9 ml) and incubated for 10 min on ice to facilitate lysis. 10 x PBS (pH 7.4, 1 ml) was added. The suspension was clarified by centrifugation (10 000 x *g*, 5 min, 4°C) and used immediately for enzyme purification using three phase partitioning.

Three phase partitioning (TPP):

No purification of *Tvi*CATL to electrophoretic homogeneity has been reported to date. So a new purification method was devised. Three phase partitioning (Pike and Dennison, 1989) was chosen as the initial purification step. Tertiary butanol [30% (v/v) of total mixture volume] and ammonium sulfate [(w/v) of the total mixture volume] were mixed with the sample solution. Tertiary butanol is infinitely miscible with water, but becomes increasingly less miscible when salt is added. As a result of this addition of ammonium sulfate to a tertiary butanol-protein mixture, proteins precipitate and a concentrated, dehydrated pellet between the salt and butanol layers is formed (Pike and Dennison, 1989). Since different proteins precipitate out at different salt concentrations, incremental addition of ammonium sulfate allows crude purification of proteins from a mixture. In the case of *Tvi*CATL, tertiary butanol and increasing percentages of ammonium sulfate [0-10%, 10-20%, 20-25% and 25-30% (w/v)] were added to optimise *Tvi*CATL isolation. The solution was centrifuged (10 000 x *g*, 10 min, RT), in a swing out rotor, after each round of ammonium sulfate addition and the protein precipitate was dissolved in Tris-HCl buffer [0.02 M, pH 7.6, 0.02% (w/v) NaN_3] before being assayed against H-D-Val-Leu-Lys-AMC for enzyme activity. The H-D-Val-Leu-Lys-AMC activity from the *T. vivax* lysate was shown to precipitate between 10 and 20% [(w/v) of the total volume] ammonium sulfate. For large scale purification of *Tvi*CATL, the final 20% (w/v) ammonium sulfate *Tvi*CATL-containing pellet was resuspended in a minimal volume of Tris-HCl buffer.

Ion exchange chromatography:

Following the initial isolation using TPP, further purification was affected using anion exchange chromatography onto a Q-Sepharose. The *Tvi*CATL-containing sample was loaded onto Q-Sepharose (26 x 110 mm, flow rate 1 cm/min, 4°C), which had been previously equilibrated with Q-Sepharose equilibration buffer [0.02 M Tris-HCl, pH

7.6, 0.02% (w/v) NaN_3,]. After elution of unbound material, as monitored by A_{280} reading returning to base line, a gradient of 0 to 1 M NaCl in Q-Sepharose equilibration buffer was applied over 5 column volumes followed by one column volume of Q-Sepharose elution buffer containing 1 M NaCl. Eluted fractions were analysed by measuring the A_{280} as well as running peak A_{280} fractions on a SDS-PAGE gel (Section 2.2.5) followed by silver staining (Section 2.2.5).

Molecular exclusion chromatography (MEC):

For further purification of *Tvi*CATL, molecular exclusion chromatography was conducted using a Sephacryl-200 HR column (25 x 650 mm, flow rate of 0.2 cm/min, 4°C). The resin has a fractionation range of 5 – 250 kDa. The column was equilibrated with one column volume of MEC buffer [0.02 M Tris-HCl, pH 7.6, 0.02% (w/v) NaN_3,]. Thereafter, the column was calibrated with 3 ml of calibration solution [2 mg/ml blue dextran (2000 kDa), 5 mg/ml of each BSA (68 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa) and myoglobin (16.7 kDa)]. The availability constant (K_{av}) was determined for each protein where the elution volume (V_e) of blue dextran denotes the void volume (V_o) and V_t is the total volumn. A Fischer's plot relating K_{av} to Log_{Mr} is presented in Fig. 2.2.

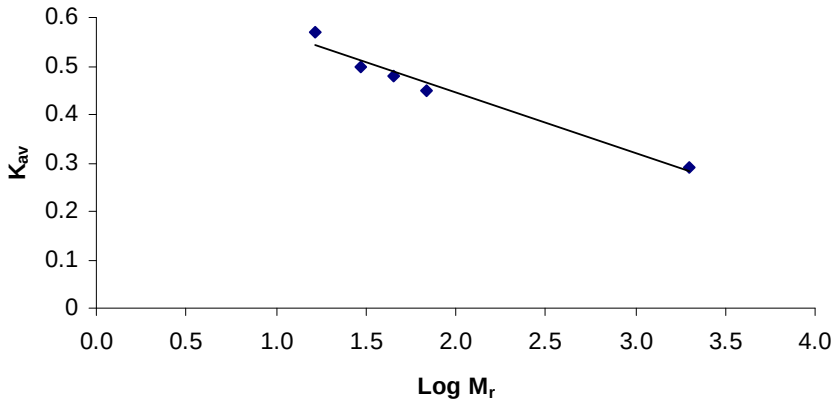


Figure 2.2 Fischer's plot for the estimation of protein Mr from MEC data. The calibration solution was applied to a Sephacryl-200 HR column (25 x 650 mm, flow rate of 0.2 cm/min, 4°C) in MEC buffer [0.02 M Tris-HCl, pH 7.6, 0.02% (w/v) NaN_3,]. The availability constant (Kav) for each protein standard was determined. The equation of the trendline is y = -7.7502x + 5.4469 with a correlation coefficient of 0.9707.

2.2.4 Quantification of *Tvi*CATL

The purified native *Tvi*CATL was quantified by using two methods, namely, BCA™ Protein Assay Kit (Pierce, Rockford, IL, USA) and the Bradford protein assay (Bradford, 1976). The BCA™ Protein Assay Kit contains a working reagent that includes bicinchoninic acid (Smith et al., 1985). The combination of Cu^{2+} to Cu^{+} reduction by protein in an alkaline medium with the highly sensitive colorimetric detection of Cu^{+} by bicinchoninic acid allows for protein quantification (Smith et al., 1985). The BCA/copper complex exhibits a strong linear absorbance at 562 nm with increasing protein concentration. The BCA colour formation is highly influenced by the amino acids cysteine, tyrosine and tryptophan. For the construction of a standard curve (Fig. 2.3), triplicate samples of known concentrations of bovine serum albumin (BSA) [0.025 – 2 mg/ml] were mixed with the working reagent (sample to working reagent ratio = 1:20) and placed in a Nunc® 96 well plate. The plate was covered and incubated at 37°C for 30 min. Colour change was measured at 562 nm using the FLUORStar Optima Spectrophotometer (BMG Labtech, Offenburg, Germany). The standard curve was used to determine the unknown concentration of purified protein.

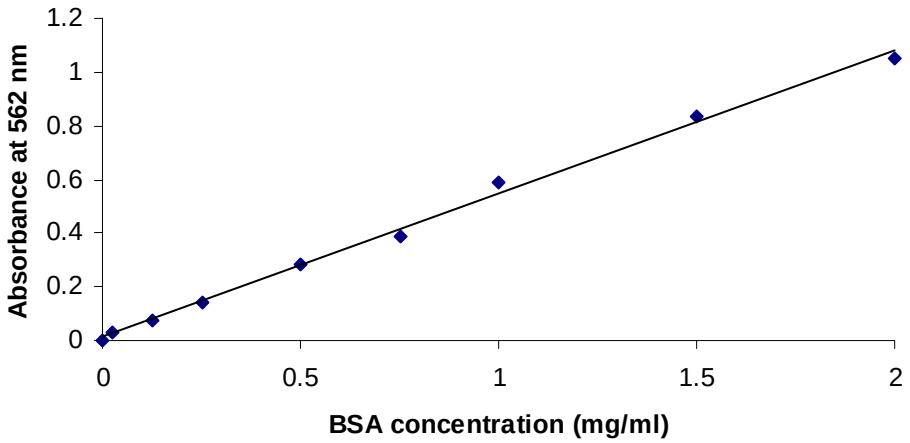


Figure 2.3 Standard curve for BCA protein quantification assay. Bovine serum albumin (BSA) samples of known concentrations (0.025 – 2 mg/ml) were measured at 562 nm in order to construct a standard curve. The equation of the trend line is y = 0.0005x + 0.0119 with a correlation coefficient 0.9968.

The Bradford protein determination method involves the binding of dye proteins, specifically arginine residues (Compton and Jones, 1985), causing a shift in the absorption maximum of the dye from 465 nm (red) to 595 nm (blue). A standard curve

was constructed using ovalbumin in a range of 1 – 5 μg (Fig. 2.4). Briefly, quintuplicate samples of 0 – 50 μl of the 100 μg/ml ovalbumin standard solution were diluted to 50 μl with dH_2O and 950 μl of Bradford dye [50 mg Serva blue G in 50 ml of 88% (v/v) phosphoric acid and 23.5 ml 99.5% (v/v) ethanol]. After colour development (2 min), the A_{595} was recorded. The standard curve was used to determine the concentration of purified protein.

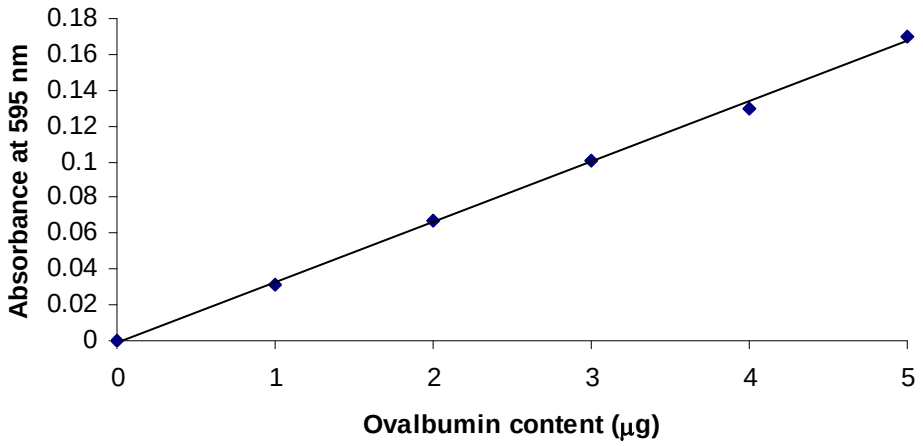


Figure 2.4 Bradford protein calibration curve. Ovalbumin solutions (1-5 μg) were used as standard solutions and their absorbance, following mixing with Bradford reagent, read at 595 nm. The equation of the trend line is given by y = 0.0338x – 0.0014, with a correlation coefficient of 0.9988.

2.2.5 SDS-PAGE gel analysis of *Tvi*CATL

The SDS-PAGE method, described by Laemmli (1970), was used to analyse protein samples. A discontinuous buffer system consisting of both a stacking buffer (500 mM Tris-HCl, pH 6.8) and a separating buffer (1.5 M Tris-HCl, pH 8.8) was used. The protein samples were treated with a reducing treatment buffer [125 mM Tris-HCl, 4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) 2-mercaptoethanol]. Disulfide bonds of protein are reduced by 2-mercaptoethanol. This allows SDS to interact efficiently with the protein to mask the charge of the protein with the negative charge of SDS (Shapiro et al., 1967). Therefore, all proteins will migrate anionically and can be separated based on differences in molecular weight when subjected to an electric field in the polyacrylamide gel (Weber and Osborn, 1969). The relative migration distance of the protein is inversely proportional to the molecular weight. Electrophoresis was conducted at 18 mA per gel for 1.5 hours in tank buffer [250 mM Tris-HCl buffer, pH

8.3, 192 mM glycine, 0.1% (w/v) SDS] using the BioRad Mini Protein III electrophoresis equipment (BioRad, USA). A calibration curve was constructed from the molecular weight markers of known size and was used to determine the molecular weight of the protein of unknown size (Fig. 2.5).

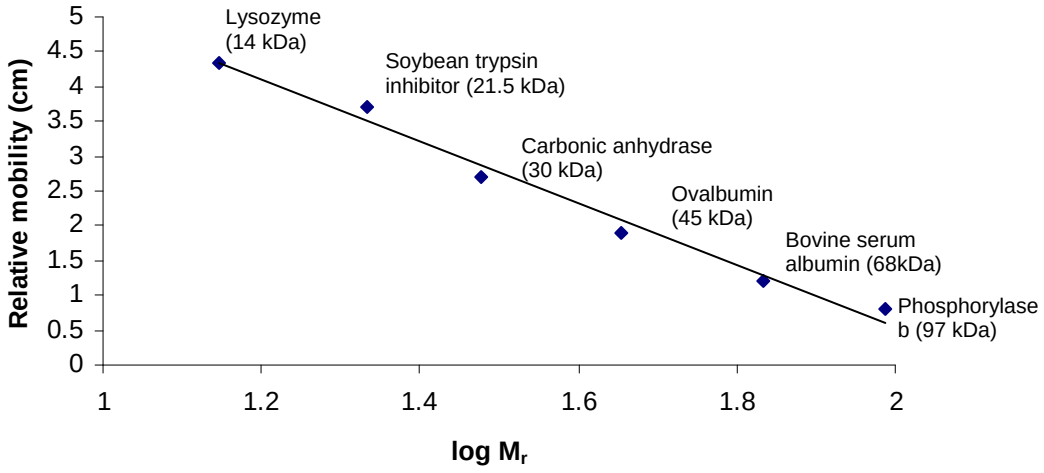


Figure 2.5 Reducing SDS-PAGE calibration curve. To determine the molecular weights of protein samples of unknown size, low range protein markers of known size were separated alongside samples to be analysed. The relative mobility of low range protein markers were measured and plotted against $\log M_r$. The equation of the trendline was y = -4.4346x + 9.4134, with a correlation coefficient of 0.9845.

To visualise electrophoresed proteins, SDS-PAGE gels were stained with Coomassie Blue [0.125% Coomassie Blue R-250, 50% (v/v) methanol, 10% (v/v) acetic acid] for 4 hours. This detection system usually has a detection limit of 50 – 100 ng of protein. For more sensitive detection, silver staining (Blum et al., 1987) was used. In this technique, following electrophoresis, the gel was placed in a meticulously clean glass dish and washed with washing solution [50% (v/v) ethanol] (3 x 20 min). The gel was subsequently soaked in pretreatment solution [0.2% (w/v) $Na_2S_2O_3.5H_2O$], which considerably reduces the intensity of background staining by preventing the precipitation of insoluble silver complexes on the surface of the gel. Following rinsing in dH_2O (3 x 20 min), the gel was soaked in impregnation solution [0.2% (w/v) $AgNO_3$, 0.028% (v/v) formaldehyde] for 20 min. The gel was rinsed again in dH_2O (3 x 20 s) and immersed in developing solution [6% (v/v) Na_2CO_3, 0.0004% (w/v) $Na_2S_2O_3.5H_2O$, 0.0185% (v/v) formaldehyde] until protein bands were visualised.

Development was stopped by soaking the gel in stopping solution [50% (v/v) methanol, 12% (v/v) acetic acid] for 10 min.

2.2.6 Western blot analysis of *Tvi*CATL

Proteins separated on a Laemmli SDS-PAGE gel were transferred from the gel onto nitrocellulose membranes using a wet blotter (BioRad Hercules, CA, USA) in transfer buffer [45 mM Tris-HCl, 173 mM glycine, 0.1% (w/v) SDS] at 40 mA (for 8 x 7 cm gels) for 16 hours according to Towbin et al (1979). Once transfer was complete, the nitrocellulose membrane was stained with Ponceau S [0.1% (w/v) Ponceau S in 1% (v/v) glacial acetic acid). When the protein bands were visible, the positions of the molecular weights markers were marked with a pencil. The nitrocellulose membrane was subsequently destained in dH_2O. After soaking the nitrocellulose in blocking solution [5% (w/v) fat-free milk powder in Tris buffered saline (TBS) (20 mM Tris-HCl buffer, pH 7.4, 200 mM NaCl)], the membrane was probed with chicken anti-C-terminal extension of *Tvi*CATL IgY diluted in 0.5% (w/v) BSA-TBS for 2 hours at RT. Following washing in TBS (3 x 5min), the membrane was incubated in rabbit anti-IgY conjugated to HRPO diluted in 0.5% (w/v) BSA-TBS. Lastly the membrane was placed in substrate solution [0.06% (w/v) 4-chloro-1-naphthol, 0.1% (v/v) methanol, 0.0015% (v/v) H_2O_2 in TBS).

2.2.7 Enzymatic characterisation

Active site titration with E-64:

The determination of the concentration of active *Tvi*CATL, using the irreversible inhibitor E-64 was done according to Barrett et al (1982). Briefly, enzyme [1 μM diluted in 0.1% (w/v) Brij-35] was incubated with E-64 (0-1 μM diluted in 0.1% Brij-35) in *Tco*CATL assay buffer (100 mM) for 30 min at 37°C. Thereafter, hydrolysis of H-D-Val-Leu-Lys-AMC (20 μM) was measured [excitation 360 nm; emission 460 nm ($Ex_{360\ nm}$; $Em_{460\ nm}$)] in a FLUORStar Optima spectrophotometer (BMG Labtech, Offenburg, Germany). The active concentration of *Tvi*CATL was determined by plotting the fluorescence against E-64 concentrations.

Gelatin SDS-PAGE analysis:

Proteolytic activity was analysed on a standard Laemmli SDS-PAGE gel, modified as described by Heussen and Dowdle (1980). A 1% (w/v) gelatin solution was prepared by dissolving in running gel buffer (1.5 ml) at 37°C. The gelatin solution was added to the remaining running gel buffer (2.25 ml) and solutions for casting a 12.5 % SDS-PAGE gel (described in Section 2.2.5). After electrophoresis, the running gel was soaked in two changes of 2.5% (v/v) Triton X-100 for an hour at room temperature to effect the removal of SDS. The gel was incubated in *Tco*CATL assay buffer [100 mM Bis-Tris buffer, pH 6.5, 4 mM EDTA, 0.02% (w/v) NaN_3, 6 mM DTT] for 3 hours at 37°C. The gel was stained with amido black solution [0.1% (w/v) amido black, 30% (v/v) methanol, 10% (v/v) acetic acid, 60% (v/v) water] for 1 hour. The gel was destained in several washes of destaining solution [30% (v/v) methanol, 10% (v/v) acetic acid, 60% (v/v) water].

Proteolysis of bovine fibrinogen:

Proteolytic activity was also tested against the protein substrate bovine fibrinogen as previously described (Rautenberg et al., 1982). Degradation at 1:100 but not 1:1000 and 1:500 molar ratios of enzyme:fibrinogen were found with native *Tbb*CATL (Troeberg, 1997) So as a starting point, a 1:100 molar ratio of enzyme:fibrinogen was used. Briefly, 0.088 μM of purified native and recombinant *Tvi*CATL (Vather, 2010) were incubated at 37°C with 0.88 μM and 8.8 μM of bovine fibrinogen respectively and removed at various time points. Proteolysis was stopped by the addition of 1 mM E-64 and samples were analysed on a 12.5% reducing SDS-PAGE gel.

Effect of pH on enzymatic activity:

For the determination of the pH-optimum for *Tvi*CATL, constant ionic strength acetate/Mes/Tris (AMT) buffers (100 mM Na-acetate, 200 mM Tris-HCl, 100 mM Mes, 8 mM DTT and 4 mM EDTA), titrated from pH 4.0 to pH 9.0, in 1.0 increments with NaOH or HCl (Ellis and Morrison, 1982) were prepared. Enzyme [1.5 ng diluted in 0.1% (w/v) Brij-35] was incubated with each buffer of a different pH for 5 min at 37°C. The substrate, H-D-Val-Leu-Lys-AMC (20 μM) was added to the solutions and the fluorescence measured ($Ex_{360\ nm}$; $Em_{460\ nm}$) FLUORStar Optima spectrophotometer (BMG Labtech, Offenburg, Germany).

2.2.8 Determination of kinetic constants

AMC standard curve:

The amount of fluorescence released during the hydrolysis of synthetic peptides was quantified by an AMC calibration curve (Fig. 2.6). Dilutions of AMC (0.5 – 15 μM) were prepared in *Tco*CATL assay buffer and reading the fluorescence ($Ex_{360\ nm}$; $Em_{460\ nm}$) in the FLUORStar Optima spectrophotometer (BMG Labtech, Offenburg, Germany).

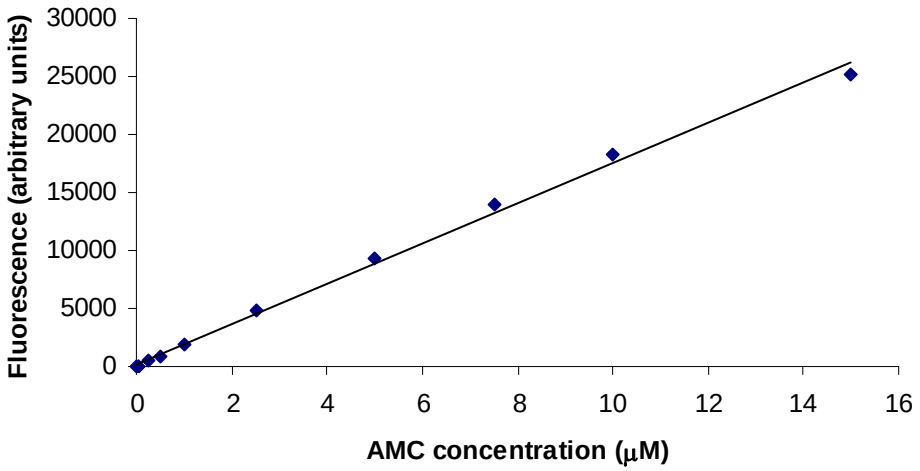


Figure 2.6 Standard curve relating amount of AMC to fluorescence ($Ex_{360\ nm}$; $Em_{460\ nm}$). The equation of the trendline is given by y = 3.4208 + 396.65, with a correlation coefficient of 0.9981.

Substrate specificity:

Substrate specificity for *Tvi*CATL was established by preincubation of enzyme as determined by E-64 titration [1.5 ng active enzyme, diluted in 0.1% (w/v) Brij-35] in *Tco*CATL assay buffer for 5 min at 37°C before substrate was added. The initial steady-state velocity (v_0) was determined by continuous assay for a range of substrate concentrations (2.5 – 50 μM) using the FLUORStar Optima spectrophotometer (BMG Labtech, Offenburg, Germany) at an excitation of 360 nm and an emission wavelength of 460 nm. The kinetic constants K_m and V_{max} were determined by hyberbolic regression of the kinetic data using Hyper32© software data (1991 – 2003, Dr. J.S. Easterby, University of Liverpool, Liverpool, UK). The k_{cat} was determined from $k_{cat} = V_{max}/[E]_0$, where $[E]_0$ represents the active enzyme concentrations (Salvesen and Nagase, 1989).

***Reversible inhibition of Tvi*CATL:**

The inhibition kinetic constant (K_i) was determined as described by Salvesen and Nagase (1989). The uninhibited rate of H-D-Val-Leu-Lys-AMC substrate hydrolysis (v_0) was first established by incubating enzyme [1.5 ng active enzyme diluted in 0.1% (w/v) Brij-35] in *Tco*CATL assay buffer for 10 min at 37°C, followed by adding 20 μM of H-D-Val-Leu-Lys-AMC and then reading the fluorescence ($Ex_{360\ nm}$; $Em_{460\ nm}$) for 5 min using the FLUORStar Optima spectrophotometer (BMG Labtech, Offenburg, Germany). Thereafter, inhibitor was added at a 20-fold molar excess over enzyme (not exceeding 5% of the total assay volume) and a new steady-state velocity determined in the presence of inhibitor (v_i). Under these conditions, the apparent inhibition constant in the presence of substrate ($K_{i(app)}$) was calculated by $v_0/v_i = 1 + [I]/K_{i(app)}$ (Salvesen and Nagase, 1989), where $[I]/K_{i(app)}$ is the slope from the plot of v_0/v_i against [I]. The true equilibrium inhibitor constant, K_i, for competitive reversible inhibitors accounts for the presence of substrate and is given by $K_i = K_{i(app)}/1 + [S]/K_m$, where [S] denotes substrate concentration.

***Irreversible inhibition of Tvi*CATL:**

The effect of irreversible inhibitors on *Tvi*CATL activity was determined under pseudo first-order conditions, with inhibitor concentration 50 times greater than the enzyme concentration (Salvesen and Nagase, 1989). *Tvi*CATL [1.5 ng active enzyme diluted in 0.1% (w/v) Brij-35] was incubated at 37°C with *Tco*CATL assay buffer and irreversible inhibitor (28 nM) to initiate inactivation. Aliquots were removed at 5 minute intervals over a for 1 hour assay period, and residual activity (v_t) determined against 20 μM H-D-Val-Leu-Lys-AMC. The pseudo first-order inhibition constant (k_{obs}) was obtained from the plot of In v_t/v_0 *versus* time where v_0 represents the initial enzymatic activity, prior to the addition of irreversible inhibitor. The apparent second-order inhibition rate constant (k_a) was obtained from the relationship $k_a = k_{obs}/[I]$, where [I] represents the inhibitor concentration. The time required for the free enzyme to decrease by 50% (half-life, $t_{1/2}$) during the reaction of the enzyme with the irreversible inhibitor is given by $t_{1/2} = 0.693/k_a[I]$ (Salvesen and Nagase, 1989).

2.3 RESULTS

2.3.1 Purification of native *Tvi*CATL

The *T. vivax* parasites were isolated from the blood of two BALB/c mice by using anion exchange chromatography on a DE-52 resin followed by the use of Histopaque® 1077, resulting in the collection of 8.5 x 10^7 parasites. The parasites were lysed and TPP was used as an initial isolation step to purify native *Tvi*CATL. TPP was optimised by adding increasing amounts of ammonium sulfate successively to parasite lysate. As it has been previously determined that the synthetic substrate H-D-Val-Leu-Lys-AMC is the preferred substrate for the recombinant catalytic domain of *Tvi*CATL (Vather, 2010), this substrate was used in an activity assay to determine the optimal percentage ammonium sulfate to affect native *Tvi*CATL precipitation (Fig. 2.7). The H-D-Val-Leu-Lys-AMC-hydrolysing activity from *T. vivax* was shown to precipitate between 10 and 20% (w/v of the total volume) ammonium sulfate. As a result the 10-20% ammonium sulphate cut was used in the TPP step of the purification procedure.

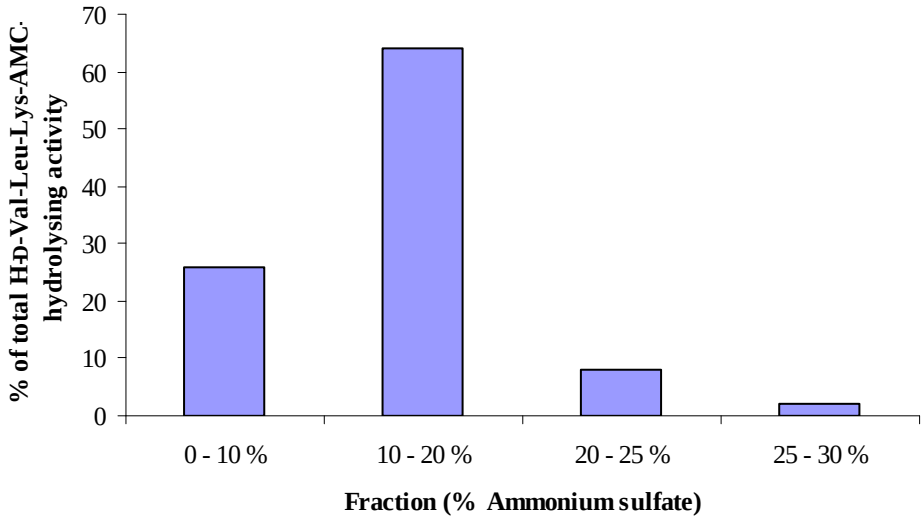


Figure 2.7 Optimisation of TPP for the isolation of *Tvi*CATL from *T. vivax*. Increasing percentages of ammonium sulfate were added to *T. vivax* lysate containing 30% (v/v of total volume) t-butanol and centrifuged. Each protein precipitate was resuspended and assayed against H-D-Val-Leu-Lys-AMC. Each reading is the average of three independent readings.

The 10-20% ammonium sulfate cut from TPP was further purified by anion exchange chromatography (Fig. 2.8). Native *Tvi*CATL eluted at 0.6 M NaCl as shown by the

single peak of absorbance at 280 nm (fractions 40 to 53, Fig. 2.8) and analysis by reducing SDS-PAGE (Fig. 2.9). Proteins from Fig. 2.9, Panel A, were further analysed by western blotting (Fig. 2.10).

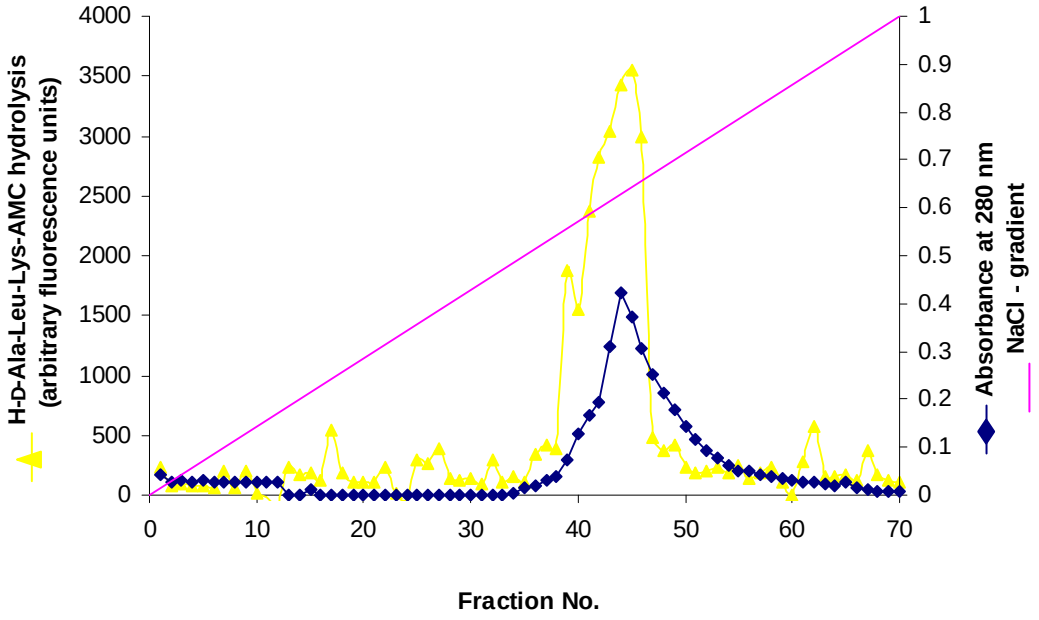


Figure 2.8 Purification of native *Tvi*CATL on a Q-Sepharose anion exchange column. The 10–20% ammonium sulfate TPP cut was loaded onto the Q-Sepharose column (26 x 110 mm, flow rate 1 ml/min,) pre-equilibrated in Q-Sepharose equilibration buffer [0.02 M Tris-HCl, pH 7.6, 0.02% (w/v) NaN_3]. Bound material was eluted with a 0–1 M NaCl gradient (-) and monitored by reading the absorbance at 280 nm (◆) and hydrolysis of H-D-Ala-Leu-Lys-AMC (▲).

Four major proteins with M_r values of 27, 56, 68 and 92 kDa were visualised (Fig. 2.9). The 56 and 68 kDa proteins constituted the bulk of the peak absorbance readings. Western blot analysis using chicken anti-*Tvi*CATL C-terminal extension peptide antibodies identified the 56 kDa protein as *Tvi*CATL.

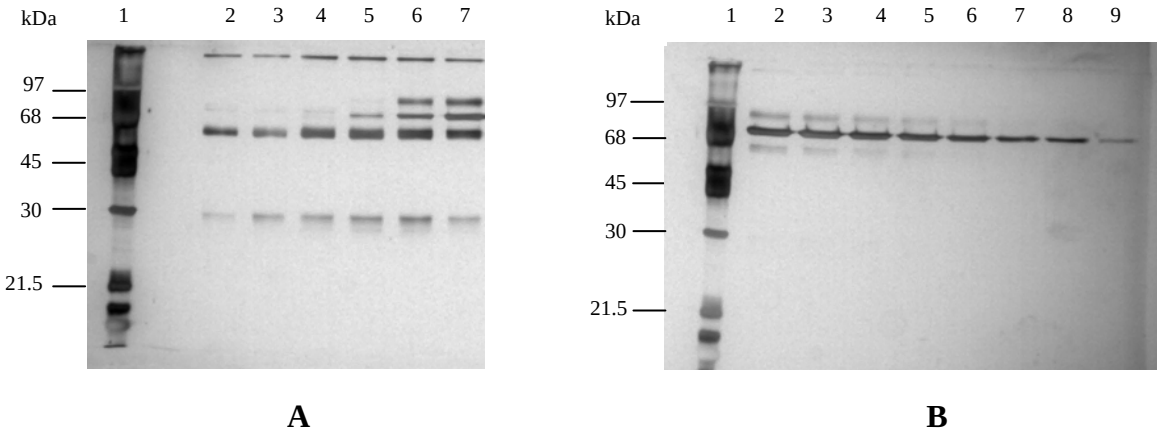


Figure 2.9 Silver stained reducing SDS-PAGE (12.5%) gel of fractions eluted from the Q-Sepharose anion exchange column. Samples (2 μg) were reduced by boiling in 2-mercaptoethanol before loading onto the SDS-PAGE gel. Panel A: lane 1, molecular weight markers; lanes 2-7, Q-Sepharose fractions 40-45. Panel B: lane 1, molecular weight markers; lanes 2-9, Q-Sepharose fractions 46-53.

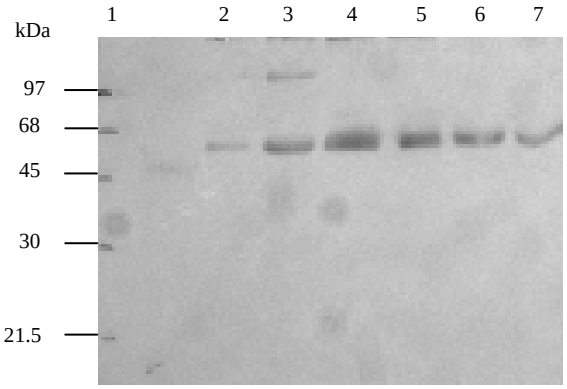


Figure 2.10 Western blot analysis of fractions eluted from Q-Sepharose. Proteins were electrophoresed on a 12.5% reducing SDS-PAGE gel and electroblotted onto a nitrocellulose membrane. The transferred protein was probed with chicken anti-*Tvi*CATL C-terminal extension peptide antibodies (5 μg/ml) followed by rabbit anti-IgY-HRPO conjugate (1:12 000). The blot was developed with substrate solution that comprised of 4-chloro-1-naphthol/H2O2. Lane 1, molecular weight markers; lanes 2-7, Q-Sepharose fractions 40-45.

Following anion exchange chromatography, fractions 40-42 (protein seen in lanes 2-4 of Fig. 2.9), were pooled, concentrated and loaded on a gel filtration chromatography column (Fig. 2.11). Only fractions 40-42 were used as previous attempts at purifying the 56 kDa *Tvi*CATL protein from fractions containing the 68 kDa protein were unsuccessful. Fractions corresponding to the A_{280} peak (Fig. 2.11) were analysed on a reducing SDS-PAGE gel (Fig. 2.11 B) and showed as a 56 kDa band but a 27 kDa contaminating protein could be observed.

As evident by the decrease in total protein levels, shown in the purification table (Table 2.1), TPP proved to be a very efficient crude purification method, removing large quantities of extraneous trypanosomal proteins from *Tvi*CATL. Anion exchange chromatography allowed for a successful step in native *Tvi*CATL isolation as indicated by an increase in specific activity.

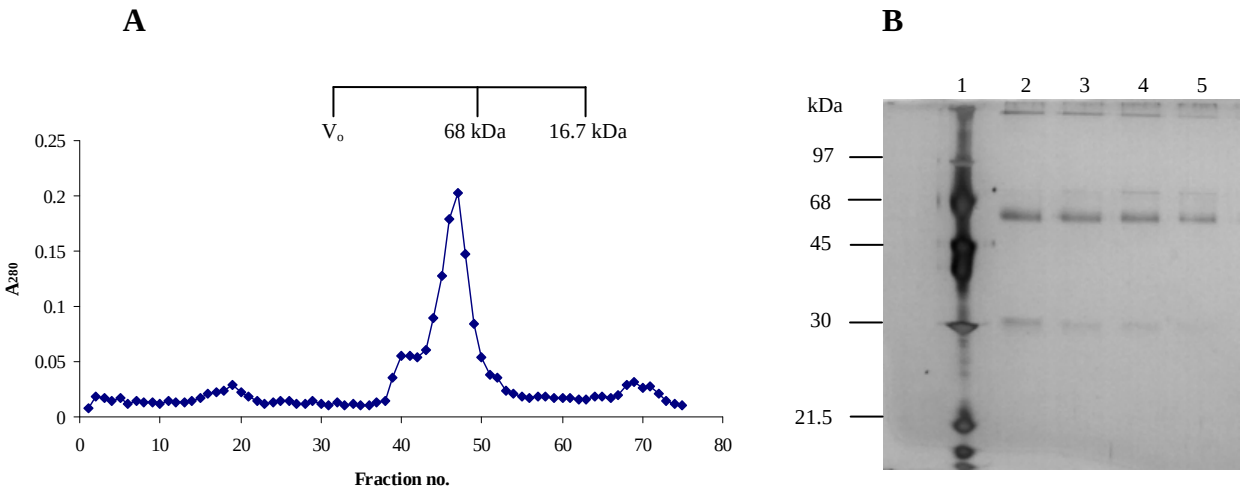


Figure 2.11 Purification of native *Tvi*CATL on a Sephacryl-200 molecular exclusion column. Panel A, elution profile on a Sephacryl-200 column (25 x 650 mm, flow rate of 1ml/min), previously equilibrated with MEC buffer [0.02 M Tris-HCl, pH 7.6, 0.02% (w/v) NaN_3,]. The column was calibrated with blue dextran (2000 kDa), bovine serum albumin (68 kDa), ovalbumin (44 kDa) and myoglobulin (16.7 kDa). Fractions collected were monitored by reading absorbance at 280 nm. Panel B, reducing 12.5% SDS-PAGE gel analysis of native *Tvi*CATL peak fractions. Lane 1, molecular weight markers; lanes 2-5 correspond to MEC peak fractions 46-49.

Table 2.1 Purification of native and recombinant *Tvi*CATL

Step	Total protein (mg)[a]	Total activity (nmol s^{-1})[b]	Specific activity (nmol s^{-1}/mg)	Purification (-fold)	Yield (%)
parasite lysate	17.65	1.90	0.11	1.00	100.00
TPP	8.19	1.36	0.17	1.55	71.58
IEC	1.47	1.18	0.80	7.27	62.10
MEC	0.60	0.51	0.85	7.73	26.84

[a] Protein concentration determined by BCA™ Protein assay kit.
[b] Hydrolysis of H-D-Val-Leu-Lys-AMC was measured in enzyme assay buffer [100 mM Bis-Tris buffer, pH 6.5, 4 mM EDTA, 8 mM DTT, 0.02% (w/v) NaN_3] at 37°C.

2.3.2 Enzymatic characterisation of native *Tvi*CATL

Active site titration of *Tvi*CATL carried out using E-64 as an active-site titrant. Native *Tvi*CATL was only ~10% active (Fig. 2.12). *Tvi*CATL did however possess gelatinolytic activity (Fig. 2.13). This activity is comparable to that of recombinant *Tvi*CATL (Vather, 2010). However *Tvi*CATL did not show any hydrolysis of the gelatin under standard conditions. Only after an extended incubation period (9 hours) was hydrolysis seen. This is consistent with the extended incubation periods of bloodstream forms of *T. vivax* lysate hydrolysing fibrinogen-containing SDS-PAGE (Mbawa et al., 1992).

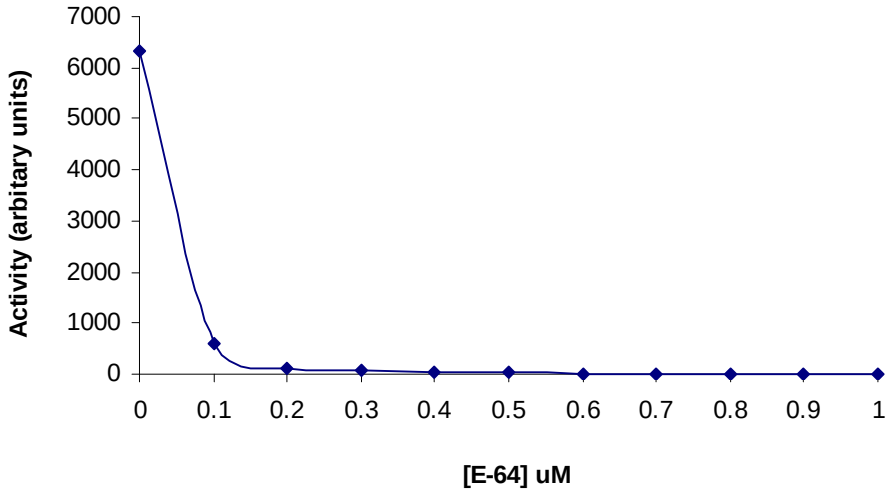


Figure 2.12 Active site-titration of native *Tvi*CATL. A constant amount of native *Tvi*CATL (1 μM) was incubated with varying concentrations of E-64 (0-1 μM) for 5 min at 37°C. Hydrolysis of H-D-Val-Leu-Lys-AMC was measured for each sample. Each data point reflects the mean activity ± SD (n = 3).

Recombinant *Tvi*CATL readily digested the α- (67 kDa) and β-chains (56 kDa) of fibrinogen after 1 hour of incubation (Fig. 2.14A). The γ-chain remained intact after the hour of incubation. Although not as definitive as the ability of recombinant *Tvi*CATL, native *Tvi*CATL was able to digest the β-chain (Fig. 2.14B). The amount of α- and γ-chain did decrease after incubation with native *Tvi*CATL, suggesting that there was a small amount of proteolysis.

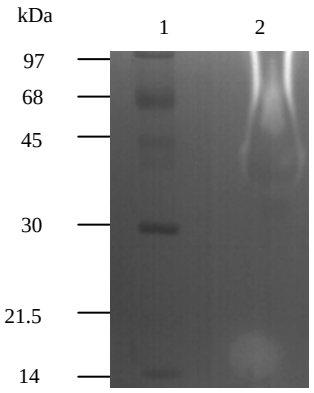


Figure 2.13 Gelatin-containing SDS-PAGE of purified native *Tvi*CATL. Non-reduced native *Tvi*CATL (20 μg, lane 2) was loaded onto a 12.5% gelatin-containing gel. Molecular masses in kDa are indicated on the left (lane 1).

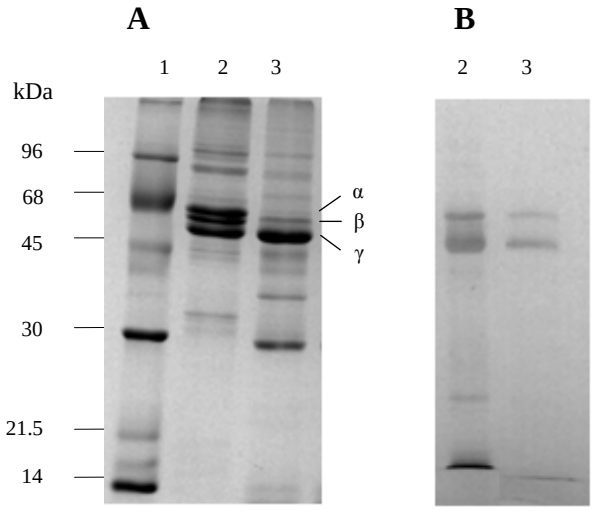


Figure 2.14 Activity of purified *Tvi*CATL against bovine fibrinogen. (A) 8.8 μM bovine fibrinogen incubated without (*lane 1*) and with (*lane 2*) 0.088 μM purified recombinant *Tvi*CATL for 1 hour at 37°C. (B) 0.88 μM bovine fibrinogen incubated without (*lane 1*) and with (*lane 2*) 0.088 μM purified native *Tvi*CATL overnight at 37°C. Proteolysis was stopped by the addition of 1 mM E-64 and samples were analysed on a reducing 12.5% SDS-PAGE gel. The molecular mass of fibrinogen α-chain is 67 kDa, the β-chain 56 kDa and the γ-chain, 47 kDa.

The pH optimum for maximal activity of *Tvi*CATL was determined using constant ionic strength AMT buffers at different pH values (Ellis and Morrison, 1982). Enzyme activity of *Tvi*CATL, at the different pH values was determined by measuring hydrolysis of H-D-Val-Leu-Lys-AMC. Native *Tvi*CATL was able to maintain some activity at very acidic and

alkaline pH values, 25% activity at pH 4.0 and 40% activity at pH 9.0 (Fig. 2.15). This pH profile is very similar to that of recombinant *Tvi*CATL. However, for recombinant *Tvi*CATL, below pH 6.0, the level of activity dropped significantly.

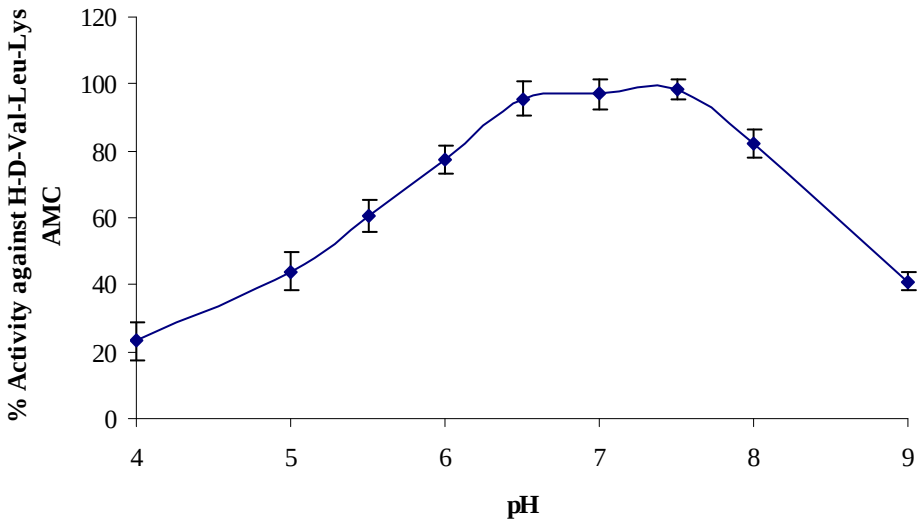


Figure 2.15 pH activity profile of native *Tvi*CATL over a pH range of 4.0-9.0. The constant ionic strength Acetate-Mes-Tris buffers were used to determine the optimum pH for maximal hydrolysis of H-D-Val-Leu-Lys-AMC by native *Tvi*CATL. Each data point reflects the mean activity ± SD (n = 3).

*Tvi*CATL hydrolysis of substrates was calculated from Lineweaver-Burk plots (Fig. 2.16). Kinetic values were tabulated in Table 2.2. For native *Tvi*CATL, hydrolysis of H-D-Ala-Leu-Lys-AMC was slightly higher than that of H-D-Val-Leu-Lys-AMC as evidenced by the higher k_{cat}/K_m value obtained for the former substrate (Table 2.2). Leu was preferred over Phe in P_2 as the hydrolysis of H-Ala-Phe-Lys-AMC was far too slow to be measured experimentally. Arg was not preferred in the P_2 position, as Z-Ala-Arg-Arg-AMC was not hydrolysed. This conclusion is supported by the fact that *Tvi*CATL was able to hydrolyse substrates with an Ala at P_3 (H-D-Ala-Leu-Lys-AMC) and an Arg in P_1 (Z-Phe-Arg-AMC). Consistent with the properties of the family C1, clan CA, *Tvi*CATL has a substrate specificity defined by the S2 pocket (Sajid and McKerrow, 2002). The nonpolar Ala and Val were both accepted in the P_3 position. This was confirmed by the fact that the k_{cat}/K_m values for H-D-Val-Leu-Lys-AMC and H-D-Ala-Leu-Lys-AMC did not differ significantly. Pro, however, was not favoured in the P_3 position as there was a 1-fold decrease in k_{cat}/K_m for H-Pro-Phe-Arg-AMC compared to Z-Phe-Arg-AMC. The fact that

activity was seen in the presence and absence of N-terminal blocked groups (Z-Phe-Arg-AMC and H-D-Val-Leu-Lys-AMC respectively), demonstrates that *Tvi*CATL possesses exo- and endopeptidase activity. *Tvi*CATL, however does not possess the typical mammalian CATB activity due to the lack of Z-Arg-Arg-AMC hydrolysis. Furthermore, aminopeptidase activity is not a characteristic of *Tvi*CATL as there was a lack of ability to hydrolyse Z-Arg-AMC.

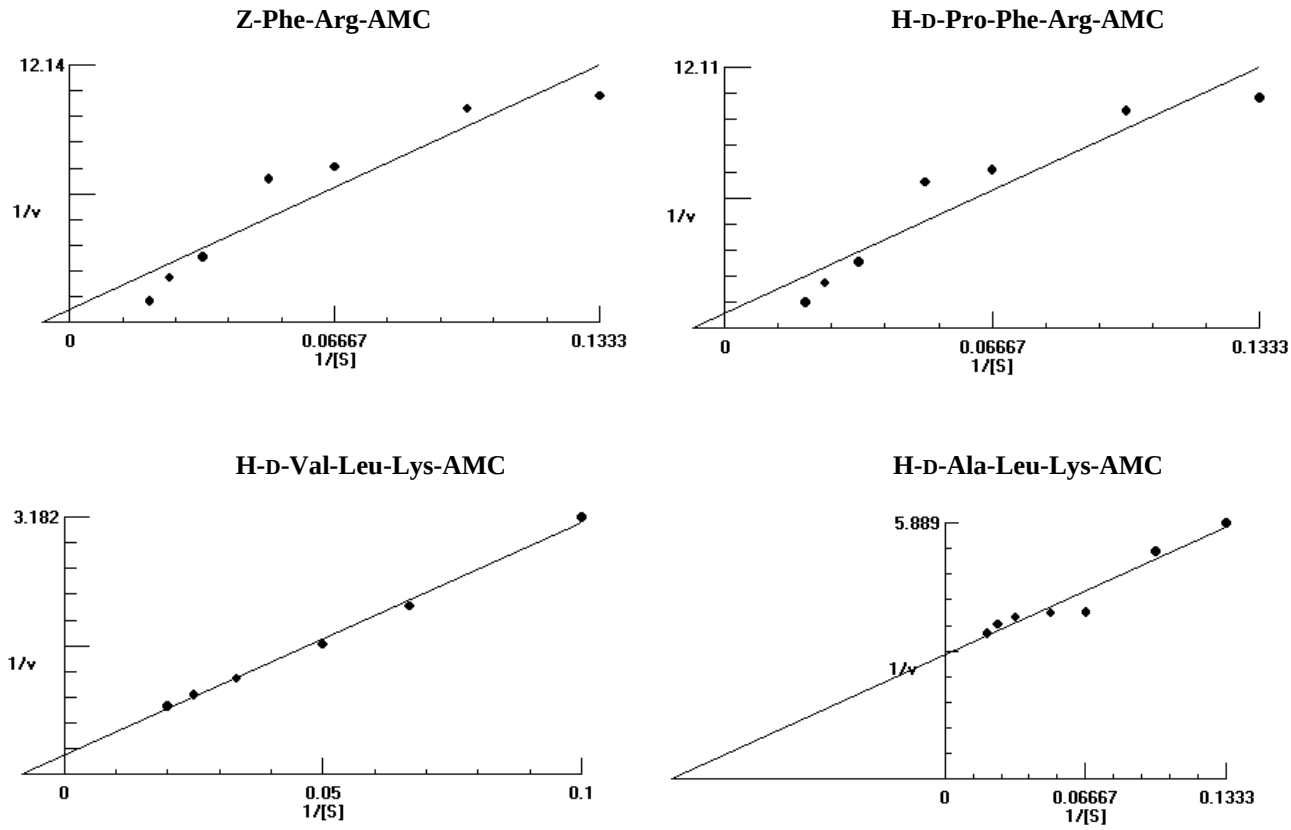


Figure 2.16 Lineweaver-Burk plot of various substrates hydrolysed by *Tvi*CATL from *T. vivax*. *Tvi*CATL was incubated with fluorogenic substrates hydrolysis of substrates was measured at Ex360 nm and Em460 nm. Lineweaver-burk plots were determined using Hyper32© software data.

Table 2.2 Kinetic parameters for the hydrolysis of fluorogenic substrates by native *Tvi*CATL[a]

Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)
Z-Phe-Arg-AMC	146.00	2.70	18.49
H-D-Val-Leu-Lys-AMC	83.70	5.04	60.21
H-D-Ala-Leu-Lys-AMC	7.74	0.56	72.42
H-Pro-Phe-Arg-AMC	128.60	1.01	7.85
H-Ala-Phe-Lys-AMC	ND[b]	ND	ND

[a] Enzyme was incubated with substrates in 100 mM Bis-Tris, pH 6.5, 4 mM EDTA, 0.02% (w/v) NaN_3, 8 mM DTT at 37°C. K_m values and the Michaelis-Menten plots were calculated using the Hyper® 32 software (J.S Easterby). [b] Not determined as reaction was too slow to be measured experimentally. No activity was detected against Z-Gly-Pro-Arg-AMC, Z-Gly-Gly-Arg-AMC, Z-Arg-AMC, Z-Arg-Arg-AMC, Z-Gly-Arg-Arg-AMC, Z-Ala-Arg-Arg-AMC and Meosuc-Asp-Tyr-Met-AMC.

A series of different competitive reversible inhibitors that can variously inhibit serine, cysteine, aspartic or metallo peptidases were tested for their ability to inhibit *Tvi*CATL (Table 2.3). The lack of inhibition by EDTA reinforces the observation that *Tvi*CATL is not metal ion dependent. Furthermore, *Tvi*CATL is not inhibited by pepstatin A, a typical aspartic peptidases inhibitor. The cysteine peptidase inhibitor, cystatin, was the most effective inhibitor with a *K*i value of 0.50 mM. This *K*i is at least 1-fold lower than that of chymostatin and antipain *K*i values. Although the results shown in Table 2.2 provides no evidence of Phe being accepted in the P_1 position, the inhibition by chymostatin suggests that Phe is accepted at such a position since chymostatin contains a Phe (attached to an aldehyde) in the P_1 position. Consistent with the lack of hydrolysis of the aminopeptidase substrate, Z-Arg-AMC (Table 2.2), bestatin, an aminopeptidase inhibitor was unable to inhibit *Tvi*CATL.

Z-Gly-Leu-Phe-CMK and H-D-Val-Phe-Lys-CMK were both able to inhibit *Tvi*CATL (Table 2.4). Consistent with its P_2 preference, Z-Gly-Leu-Phe-CMK, with Leu in P_2, was able to inhibit *Tvi*CATL better than H-D-Val-Phe-Lys-CMK. No inhibition was noted for typical serine peptidase inhibitors, PMSF, TLCK and TPCK. The alkylating thiol group

inhibitors, iodoacetamide and iodoacetic acid successfully inhibited *Tvi*CATL. No significant difference in $t_{1/2}$ and k_a was seen between these inhibitors.

Table 2.3 Competitive reversible inhibitors of native *Tvi*CATL[a]

Inhibitor[b]	*K*i (mM) (x10^3)
Leupeptin	1.24
Chymostatin	4.50
Antipain	15.05
CEW[c] cystatin	0.50

[a] Enzymes were incubated with inhibitors in 100 mM Bis-Tris, pH 6.5, 4 mM EDTA, 0.02% (w/v) NaN_3, 8 mM DTT at 37°C.
[b] No inhibition was observed with bestatin (10 μM), EDTA (10 mM) or pepstatin A (1 μM).
[c] Chicken egg white

Table 2.4 Irreversible inhibitors of native *Tvi*CATL [a]

Inhibitor [b]	k_a ($mM^{-1}.s^{-1}$)[c]	$t_{1/2}$ (s)
Z-Gly-Leu-Phe-CMK	89.50 ± 11.21	254
Iodoacetamide	35 ± 2.76	652
Iodoacetic acid	31.50 ± 1.01	724
H-D-Val-Phe-Lys-CMK	74.02 ± 6.23	321

[a] Enzymes were incubated with inhibitors in 100 mM Bis-Tris, pH 6.5, 4 mM EDTA, 0.02% (w/v) NaN_3, 8 mM DTT at 37°C. No inhibition was detected with Boc-VLGK-CHN_2-HoAC, PMSF, phenylmethanesulfonyl fluoride; TLCK, Nα-Tosyl-Lys-chloromethyl ketone and TPCK, L-1-tosylamido-2-phenylethyl chloromethyl ketone.
[b] 28 nM of inhibitor used.
[c] Data reflect the mean k_a ± S.D. (n=3).

2.4 DISCUSSION

Since drug resistance in *T. congolense* has been reported in 17 African countries of sub-Saharan Africa (Delespaux et al., 2008) and the African agriculture loss due to trypanosomosis is US $ 4.75 billion per year (Van den Bossche et al., 2010) show that new measures need to be put in place to control the disease. The ability of *Tco*CATL to degrade

a range of proteins *in vitro* (Mbawa et al., 1992) and the fact that chloromethylketones, fluoromethylketones and diazomethylketones kill bloodstream forms of *T. b. brucei in vitro* (Troeberg et al., 1999) and some peptidyldiazomethylketones (Z-Leu-Leu-Met-CHN_2, Z-Leu-Met-CHN_2 and Z-Leu-Lys-CHN_2) are trypanocidal against *T. congolense in vitro* (Mbawa et al., 1992) provide promising possibilities for drug and vaccine development. Furthermore, preclinical proof-of-concept studies of K11777 with *Tcr*CATL provide hope for chemotherapeutics (Kerr et al., 2009).

The results presented here show that a cysteine peptidase, *Tvi*CATL, has been purified and characterised for the first time from *T. vivax* strain Y486. The initial growth of parasites in BALB/c mice presented an alternative to harvesting parasites grown in rats. Rodent models have been shown to be resistant to *T. vivax* infection (Terry, 1957) and as a result, would not have given a high enough level of parasitaemia to obtain sufficient numbers of parasites for enzyme isolates. A satisfactory yield of native *Tvi*CATL was purified from the parasites. Likewise, native *Tco*CATL yielded 24% enzyme (Mbawa et al., 1992). However, the yield for native *Tbb*CATL, 5.3%, was less (Troeberg et al., 1996). The initial purification step used for native *Tvi*CATL was TPP, removing more than half of the extraneous trypanosomal total protein content. As has previously been described by Pike and Dennison (1989), this technique is cheap, effective and easy to perform. The chromatography procedures, ion exchange chromatography and molecular exclusion chromatography for native *Tvi*CATL proved to be successful final purification methods, yielding pure protein. As seen from the total protein, *Tvi*CATL constitutes as much as 3.4 % of the total *T. vivax* protein content. Similarly, as much as 0.5-1.5% of the total *T. congolense* protein content may be cysteine peptidases as was purified by thiopropyl-Sepharose affinity chromatography followed by molecular exclusion chromatography (Mbawa et al., 1992).

Similar to the recombinant catalytic domain of *Tvi*CATL (Vather, 2010), native *Tvi*CATL possesses most of its activity around neutral pH. The only variation between native and recombinant catalytic domain *Tvi*CATL (Vather, 2010) pH profile, is the ability of the recombinant catalytic domain of *Tvi*CATL to remain active at acidic pH values. Whether or not the C-terminal extension of *Tvi*CATL influences the activity at acidic pH values values is still to be determined. It has, however, been shown that full length *Tbr*CATL as

well as truncated *Tbr*CATL (without the C-terminal extension) have very similar pH profiles (Caffrey et al., 2001). Native *Tcr*CATL (Chagas et al., 1997), *Tco*CATL (Chagas et al., 1997; Mbawa et al., 1992) and *Tbb*CATL (Troeberg et al., 1996) have pH optima slightly more acidic than native *Tvi*CATL.

*Tvi*CATL efficiently hydrolysed H-D-Val-Leu-Lys-AMC and H-D-Ala-Leu-Lys-AMC. However, the inability to hydrolyse H-Ala-Phe-Lys-AMC demonstrates that Leu is favoured over Phe in P_2. Further support for such slight preference was deduced from the lower k_{cat}/K_m values obtained for H-Pro-Phe-Arg-AMC and Z-Phe-Arg-AMC compared to the substrate with Leu in P_2. It was concluded that Leu was preferred for the P_2 position and Lys for the P_1 position. Similar trends have been seen for both *Tcr*CATL and *Tbr*CATL, studied using a completely diversified positional scanning synthetic combinatorial library**,** where fluorogenic peptide substrates were used that allowed for the construction of substrate libraries to rapidly identify the primary and extended specificity of peptidases (Choe et al., 2006; Harris et al., 2000). Experimentally, recombinant *Tcr*CATL and *Tbr*CATL showed a marked preference for a hydrophobic residue at P_2 with Leu being favoured. Both native *Tco*CATL (Mbawa et al., 1992) and *Tbb*CATL (Troeberg et al., 1996) rapidly hydrolyse substrates with basic residues such as Arg or Lys in P_1 and hydrophobic residues such as Phe and Leu in P_2. However there was a slight preference for Phe in the P_2 position in both cases. It has also been recently reported that *Tco*CATL has a clear preference for Leu in P_2 and Lys in the P_1 position (Pillay et al., 2010).

It has been previously shown, recombinant *Tcr*CATL2 (a cysteine peptidase isoform) has a marked preference for Leu over Phe (dos Reis et al., 2006), whereas recombinant *Tcr*CATL accepts Phe and Leu in the P_2 position with only a slight preference for Leu (Lima et al., 2001). This suggests that the S_2 pocket of *Tcr*CATL2 might be more restricted than that of *Tcr*CATL. Structural studies of *Tcr*CATL have shown that, when in complex with irreversible inhibitors, Leu189, Ala258 and Leu282 (*Tcr*CATL numbering as in Fig. 2.1) line the S_2 pocket and readily accommodate the Phe residue at P_2, which interacts with the enzyme's backbone via the carbonyl of Gly188 (Brinen et al., 2000). However, as was reported that the Ser258 substitution found in recombinant *Tcr*CATL2, combined with the proximal change of Met190 to Ser190 may ultimately influence the enzyme's S_2 specificity (dos Reis et al., 2006). In the case of *Tvi*CATL, residues that also supposedly line the S_2

pocket (based on the sequence alignment in Fig. 2.1) are similar to *Tcr*CATL and do not possess the changes, found for *Tcr*CATL2, in the lining of the S_2 pocket. This supports the observation that *Tvi*CATL has only a slight preference for Leu over Phe in the S_2 position. Recent data have shown that a novel cysteine peptidase from *Plasmodium vivax*, VX-4, along with two other cysteine peptidases from *P. vivax*, VX-2 and VX-3 prefer Leu over Phe in the P_2 position at pH 5.5 and pH 6.5 (Na et al., 2010). It is of interest that preference of Leu over Phe in P_2 is a feature observed in mammalian CATS and CATK, but not with CATL or CATB (Choe et al., 2006). The inability of both native and recombinant *Tvi*CATL to hydrolyse Z-Arg-Arg-AMC demonstrates that lack of a CATB-like character. However, native and recombinant *Tcr*CATL have the ability to hydrolyze substrates containing the positively charged residue, Arg, in P_2 (Judice et al., 2001; Lima et al., 1992; Nery et al., 1997). This has been attributed to the presence of the Glu 333 (green arrow in Fig. 2.1) residue at the bottom of the S_2 pocket, which interacts with the positive charge, stabilising basic residues in the pocket (Gillmor et al., 1997). In the case of *Tco*CATL and *Tvi*CATL, as shown in the sequence alignment (Fig. 2.1), a Leu residue instead of a Glu residue occupies this position which suggests that there would be a difficulty in accommodating Arg in the P_2 position. For TbrCATL, an Ala residue occupies this position and has been previously shown that it too is unable to degrade substrates with an Arg in P_2 (Caffrey et al., 2001).

Native *Tvi*CATL was effectively inhibited by chicken egg white cystatin, similar to that of native *Tco*CATL (Mbawa et al., 1992) and *Tcr*CATL (Lima et al., 2001). It was previously speculated that because of conserved sequences amongst cystatins and the highly similar sequences of trypanosomal cysteine peptidases to those of the mammalian hosts, trypanosomal cysteine peptidases are likely to be inhibited by cystatins in the bloodstream of mammalian hosts (Troeberg et al., 1996). The peptidyl-diazomethane inhibitor based on a conserved sequence of cystatins, Z-RLVG-CHN_2, inhibits native *Tco*CATL and *Tcr*CATL (Chagas et al., 1997), however the mammalian homologs such as CATB and CATL are inhibited to a greater extent with such peptidyl-diazomethane inhibitors. Although no such cystatin-conserved peptidyl-diazomethane inhibitor was tested for *Tvi*CATL, it can be considered that such an inhibitor would be able to effectively inhibit *Tvi*CATL, based on the similarity in P_2 specificity of *Tvi*CATL and mammalian CATL.

The hydrophobic nature of the P_2-binding subsite was further confirmed by the ability of both Z-Gly-Leu-Phe-CMK and H-D-Val-Phe-Lys-CMK, to inhibit native *Tvi*CATL. Native *Tco*CATL is inhibited by Z-Leu-Met-CHN_2 and Z-Leu-Lys-CHN_2, with hydrophobic groups in P_2 (Leu) and either a hydrophobic or basic residue in the P_1-binding subsite (Mbawa et al., 1992). Although not tested in the present study, vinyl sulfones have been proven to have desirable attributes including selectivity and stable inactivation of the target enzyme (Palmer et al., 1995). Both vinyl sulfones, K11017 (Mu-Leu-Pph-VSPh) and K11777 (N-Mpip-Phe-Hph-VSPh) have been tested with recombinant falcipain-3 (cysteine peptidase from *Plasmodium falciparum*), *Tcr*CATL and *Tbr*CATL (Kerr et al., 2009). Both vinyl sulfones have a homophenylalanine in the P_1 position with variation at P_2; Leu for K11017 and Phe for K11777. K11017 has also been tested against native *Tbb*CATL and was one of the most effective tested inhibitors (Troeberg et al., 1999). It can be suggested that such a vinyl sulfone will be able to effectively inhibit *Tvi*CATL as Z-Gly-Leu-Phe-CMK, with a Leu in P_2 and a Phe in P_1 was an efficient inhibitor. The homophenylalanine group, with its extra branch was chosen for such synthetic inhibitors as it was previously observed that the incorporation of the homophenylalanine side chain at P_1 resulted in inhibitors that were not degraded by human peptidases, in contrast to natural amino acids at this position (Huang et al., 2003).

Clan CA peptidases are characterised by their sensitivity to the general cysteine peptidase inhibitor E-64 and have a substrate specificity defined by the S_2 pocket (Sajid and McKerrow, 2002). Our data indicate that such characterisation of both native and recombinant *Tvi*CATL is consistent with that of Clan CA peptidases. Whether or not *Tvi*CATL, like the *T. congolense* counterpart, *Tco*CATL, acts as a pathogenic factor, is still to be determined. Further structural and functional characterisation of *Tvi*CATL in *T. vivax* will be useful in facilitating the understanding of such a cysteine peptidase.

CHAPTER 3
CRYSTALLISATION OF TRYPANOSOMAL CYSTEINE PEPTIDASES

3.1 INTRODUCTION

African animal trypanosomosis (Nagana) is a widespread disease, affects 40 African countries and is costing the African continent US $4.75 billion per year (Van den Bossche et al., 2010). *T. congolense* along with *T. vivax* and *T. brucei brucei* are the three main trypanosomal species that infect livestock with the former two species the most pathogenic, causing Nagana. There are three main phases of bovine trypanosomosis. These include the early acute phase, stabilisation and chronic phase (Taylor and Authié, 2004). Acute phase is usually accompanied by a fever and anaemia. Cachexia, impairment of milk productivity and oestrous cycle are signs of the chronic phase. The chronic phase can lead to death, self-cure or to a persistent carrier state.

West-African taurine breeds of cattle have a reduced susceptibility to the disease, trypanosomosis. The reduced susceptibility is commonly referred to as trypanotolerance and is defined as the trait that confers the capacity to survive and remain productive during and after trypanosome infection (Murray et al., 1982). Trypanotolerant cattle have reduced mortality, trypanosome levels, anaemia and furthermore, reproductive performance is better than indicine breeds such as Boran cattle (*Bos indicus*) (Naessens, 2006). Trypanotolerant cattle are able to mount an antibody response to the cysteine peptidase, *Tco*CATL (Lalmanach et al., 2002) better than trypanosusceptible cattle breeds.

The trypanocidal compounds such as isometamidium chloride, homidium (bromide and chloride) and diaminazene are used for primary control of the disease (Holmes et al., 2004). However these drugs have been used for more than 50 years and as a result drug resistance has developed (Kroubi et al., 2011). Other major strategies for controlling the disease include eradication or control of the tsetses fly vector. Vaccination strategies have been hampered by the ability of trypanosomes to continuously evade the host antibody response with the generation of an inexhaustible collection of variable surface

glycoproteins (VSGs). Non-variable antigen-based and anti-disease vaccines (Authié, 1994) have been seen as alternative approaches. For the anti-disease vaccine approach, the artificial induction of antibodies against *Tco*CATL could render a disease-susceptible host more tolerant (Authié et al., 1992; Lalmanach et al., 2002).

Like mammalian CATL, *Tco*CATL and other trypanosomal cysteine peptidases hydrolyse substrates using the catalytic triad, Cys25, His159, Asn175 (papain numbering). Trypanosomal CATL-like cysteine peptidases, like the mammalian homologs, are synthesised as pre-pro-proteins (Cazzulo et al., 1989; Eakin et al., 1992; Sakanari et al., 1997). However, unlike their mammalian counterparts, an additional 11-13 kDa C-terminal extension, of unknown function, is separated from the catalytic domain by poly-threonine (*Tcr*CATL and *Tvi*CATL) or poly-proline (*Tbb*CATL and *Tco*CATL) sequences.

Trypanosomal cysteine peptidases offer good targets for chemotherapy. A large amount of evidence supports this. *Tbb*CATL has been linked to the ability of *T. b. brucei* to cross the blood brain barrier (Abdulla et al., 2008). Amastigote replication, intracellular development and host cell invasion have been implicated with *Tcr*CATL function. (Aparicio et al., 2004). Besides the link to trypanotolerance, *Tco*CATL degrades a range of proteins *in vitro* (Mbawa et al., 1992). Like *Tbb*CATL and *Tcr*CATL, *Tco*CATL is active at neutral pH allowing it to interact with host proteins in a physiological environment (Lalmanach et al., 2002). Incomplete trypanosomal cysteine peptidase inhibition with host cysteine peptidase inhibitors has been documented (Lonsdale-Eccles et al., 1995; Wiser et al., 1997). In some cases, biologically active molecules have been released from such inhibition and as a result may play a role in pathology. Pyroglutamyl peptidase I (PGP), which has a slightly different catalytic triad (Glu instead of Asn), is entirely responsible for the reduced plasma half-life of thyrothrophin releasing hormone and partially responsible for the reduced half-life of gonadothrophin releasing hormone in *T. b. brucei* infected rodents (Morty et al., 2006). *T. congolens*e PGP has similar enzyme activity to *T. b. brucei* PGP (Mucache, 2011). The cysteine peptidase inhibitor, E-64, which binds to the thiol group of the active-site Cys residue, capable of inhibiting *Tbb*CATL (Troeberg et al., 1996), *Tcr*CATL (Lima et al., 2001), *Tco*CATL (Mbawa et al., 1992) and *Tvi*CATL (Chapter 2 of this dissertation), does not inhibit *T. congolense* PGP (Mucache, 2011). The three-dimensional structures of the peptidases are thus required to explain this difference in susceptibility to E-64 inhibition.

Inhibitors, such as the irreversible dipeptidyl substrate analogues coupled to reactive fluoromethylketone (CH_2F) or diazomethylketone (CHN_2) groups have been widely researched for selectively inhibiting the parasitic cysteine peptidases. Some peptidyldiazomethylketones (Z-Leu-Leu-Met-CHN_2, Z-Leu-Met-CHN_2 and Z-Leu-Lys-CHN_2) were trypanocidal when tested against bloodstream forms of *T. congolense* (Mbawa et al., 1992) and *T. b. brucei in vitro* (Troeberg et al., 1999). Peptidyl fluoromethylketones limited the survival of *T. cruzi* trypomastigotes (Ashall et al., 1990). A proposed binding mode for these peptidyldiazomethylketones have shown that the peptide component is bound by specificity subsites S_1, S_2 and S_3 (Barrett et al., 1982). The availability of crystal structures has allowed for the directed design of synthetic peptidase inhibitors (Chen et al., 2008). The peptidomimetic vinyl sulfone, K11777 (N-Mpip-Phe-Hph-VSPh), which targets *Tcr*CATL, is lethal to *T. cruzi* parasites both *in vitro* and in animal models of the disease (Engel et al., 1998a; Engel et al., 1998b) and is an example of 'chemical knockout' of the *Tcr*CATL gene. The crystal structure of recombinant *Tcr*CATL in complex with K11017 (Mu-Leu-Hph-VSPh) shows that the hydrophobic residue in the P_2 position is able to make contact with the S_2 subsite of *Tcr*CATL as the S_2 subsite is lined with non-polar residues (Bryant et al., 2009; Kerr et al., 2009). Furthermore, the phenyl vinyl sulphone in the in the P_1' position is able to rest on the floor of the S_1' subsite and covalently bind with the active-site Cys residue (Chen et al., 2008). The Mu rather than the N-Mpip group in the P_3 position is better accommodated by *Tbr*CATL, from *T. b. rhodesiense*, and *Tcr*CATL as it is able to make more polar interactions through bridging water molecules (Kerr et al., 2009). The K11777, which has no significant toxicity to rodent or canine hosts (Abdulla et al., 2007), is currently completing pre-clinical testing for the treatment of Chaga's disease.

To date the structure of *Tco*CATL has not been elucidated. The only two trypanosomal cysteine peptidase structures determined are those of *Tcr*CATL (Eakin et al., 1993; McGrath et al., 1995) and *Tbr*CATL (Kerr et al., 2009). *Tco*CATL shares 68% sequence identity with the central catalytic domain of *Tcr*CATL (Lalmanach et al., 2002). The predicted 3D structure of *Tco*CATL, based on the sequence homology with *Tcr*CATL, follows a common framework found for papain-like cysteine peptidases (Fig. 1.6, Chapter 1 of this dissertation). The framework consists of two domains: the L domain, which is mainly α helical; and the R domain, which is mostly composed of antiparallel β sheets

(Lecaille et al., 2001) (Fig. 1.11, Chapter 1 of this dissertation). The catalytic triad and the residues defining the extended substrate-binding site are located in the cleft between the R domain and the L domain (Lecaille et al., 2001). *Tco*CATL, based on the hydrophobic pocket of the S_2 subsite, has a more restricted specificity than *Tcr*CATL (Chagas et al., 1997) and was previously confirmed using the superimposition model of *Tco*CATL and *Tcr*CATL (Lecaille et al., 2001). However, unlike *Tcr*CATL, *Tco*CATL is monomeric at acidic pH and dimeric at neutral pH (Boulangé et al., 2011). Antibodies directed against *Tco*CATL are better at recognising the dimeric form (Boulangé et al., 2011). It is therefore crucial to understand the dimerisation mechanism of *Tco*CATL. Some of the amino acid residues, that have been shown to be involved in dimerisation through a salt bridge, were shown through site-directed mutagenesis to be Lys39, His43 and Glu44 (Ndlovu, 2010). At physiological pH the His43 is deprotonated resulting in a salt bridge between Lys39 and Glu44, thus forming a β-turn, giving the protein dimeric conformation (Ndlovu, 2010). Mutating Lys39 and Glu44 to Phe39 and Pro44, prevents salt bridge formation resulting in a monomeric *Tco*CATL conformation. Furthermore, as previously speculated, mutating His43 to Trp43 reults in the protein existing as a dimer (Ndlovu, 2010).

Structural genomics initiatives have demonstrated that for a given protein construct, the initial likelihood of obtaining a crystal is between 30%-40% (Rupp, 2009) and the probability of a crystal to yield usable diffraction data again only ranges between 30%-40%. That is why it is imperative that the crystallisation strategy be carefully selected and that careful observation and interpretation is carried out to increase the chances of crystallisation. Various strategies can be employed which include commercially available crystallisation screens, which screens a wide variety of conditions, or limiting screening based on information collected from proteins of similar sequence and character that have been crystallised (Fig.1.10, Chapter 1 of this dissertation). Generally, commercially available screens are used to maximise the chances of producing well-diffracting crystals initially (Rupp, 2009). If positive results are found from initial screening, optimisation screens can be done. These usually include exploring the phase diagram (Fig.1.9, Chapter 1 of this dissertation), or level of supersaturation, for the protein in a certain crystallising reagent (Benvenuti and Mangani, 2007) and adding various additives such as detergents, reducing agents and heavy metal compounds to the screens. If after several optimisation screens no positive results are obtained, protein construct manipulation presents a further

avenue to explore in order to obtain crystals. Non-diffracting crystals also require further optimisation by either mutation of amino acid residues, changing the protein construct and/or screening of the protein with other crystallising reagents and buffers.

In the present study, the catalytic domain of *Tco*CATL and *Tvi*CATL, as well as the full-length pyroglutamyl peptidase, PGP, were recombinantly expressed and purified. These proteins along with *Tco*CATL dimerisation mutants, *Tco*CATL (H43W), *Tco*CATL (K39F; E44P) and the proenzyme mutant *Tco*CATL (C25A), were screened for suitable crystallisation conditions using the recently acquired Rigaku robotic crystallisation suite (Rigaku Automation Inc., USA), to establish the technology at our Institution. Crystallisation hits for *Tco*CATL (K39F; E44P) were further optimised and their X-ray diffracting ability evaluated.

3.2 MATERIALS AND METHODS

3.2.1 Materials

***P. pastoris expression and purification*:**

Glycerol stocks of *P. pastoris* GS115 cells containing recombinant pPic9-*Tco*CATL and pPic9-*Tvi*CATL were obtained from Dr A Boulangé (University of KwaZulu-Natal, Pietermaritzburg) and Professor Theo Baltz (University Victor Segalen, Bordeaux, France) respectively. Ten mg/ml of purified and concentrated *Tco*CATL (H43W), *Tco*CATL (C25A) and *Tco*CATL (K39F; E44P) were supplied by Hlumani Ndlovu (Ndlovu, 2010). SP Sephadex C-25, Sephacryl S-200, 10 kDa cut-off Snakeskin® pleated dialysis tubing, tetracycline, ampicillin and chloramphenicol were purchased from Sigma (Munich, Germany). The VersaDoc imaging system using Quantity One software was purchased from Bio-Rad (USA). Ultrospec 2100 *pro* spectrophotometer was from Amersham Biosciences (Sweden). Whatman No. 1 filter paper was purchased from Whatman (England). Centriprep ultrafiltration devices with YM-10 membrane were purchased from Millipore (MA, USA).

E. coli expression and purification:

Glycerol stocks of *E. coli* BL21 (DE3) containing pGEX4T1-PGP were obtained from Mucache (2011). Ampicillin and reduced glutathione were purchased from Sigma (Munich, Germany). GST-bind column was purchased from Sigma (Munich,

Germany).Thrombin and a thrombin cleavage buffer and were purchased from Novagen (WI, USA). Poly-Prep® chromatography columns were from Biorad (USA).

Antibodies:

Anti-peptide antibodies were produced in chickens against a peptide corresponding to the 22 N-terminal residues of *Tco*CATL (Mkhize, 2003). Antibodies against the recombinant C25A mutant full length inactive form of *Tvi*CATL (ΔFL*Tvi*CATL) were raised in chickens (Vather, 2010). The rabbit anti-IgY-HRPO conjugate was obtained from Sigma (Munich, Germany). Nunc-Immuno™ 96-well plates were purchased from NuncIntermed (Denmark).

Enzyme assays:

The peptide substrate, Z-Phe-Arg-AMC, pGlu-AMC and the inhibitor E-64 were purchased from Sigma (Munich, Germany). Nunc Black 96-well plates were purchased from NuncIntermed (Denmark). The plates were read in a FLUORStar Optima Spectrophotometer (BMG Labtech, Offenburg, Germany).

Crystallisation:

Joint Centre for Structural Genomics (JSCG) + and the pH, anion and cation testing (PACT) commercial screens were purchased from Molecular Dimensions (United Kingdom). The Alchemist II, CrystalTrak™ software, UV DesktopMinstrel, Phoenix software and Gallery 160 Plate Hotel were all purchased from Rigaku Automation, Inc. (USA). Sitting-drop crystallisation plates (96 wells), deep-well blocks (96 wells) and UV crystallisation plate sealants were all purchased from Molecular Dimensions (United Kingdom). Capillary tubes were from Marienfeld-Superior (Germany). Crystals in capillary tubes were viewed using an Olympus CKX41 Inverted Microscope (Wirsam, South Africa). Cell culture plates (24-wells) were purchased from SPL Lifesciences (Korea) and microscope coverslips were from Lasec (South Africa).

3.2.2 Expression of the recombinant catalytic domains of *Tvi*CATL and *Tco*CATL in *P. pastoris*

The glycerol stocks of *P. pastoris* GS115 cells containing pPic9-*Tco*CATL and pPic9-*Tvi*CATL were streaked on YPD plates [1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose, 15 g/l bacteriological agar] containing ampicillin (50 µg/ml). The plates were incubated at 30ºC for 3 days. Liquid YPD (100 ml) containing ampicillin (50 µg/ml) was inoculated with a single colony of *P. pastoris* yeast. The culture was incubated at 30ºC in a baffled flask on an orbital shaking incubator for 2 days. The culture was transferred into 1 L BMGY [1% (w/v) yeast extract, 2% (w/v) peptone, 100 mM potassium phosphate buffer, pH 6.5, containing 1.34% (w/v) yeast nitrogen base without amino acids (YNB)]. The culture was grown at 30ºC for a further 2 days (until OD_{600} of 4-6 was reached) to increase biomass. Cells were pelleted (6000×g, 10 min, RT) and resuspended in 500 ml buffered minimal medium [BMM, 100 mM potassium phosphate buffer, pH 6.5 containing 1.34% (w/v) YNB, 0.0004% (w/v) biotin, 5% (v/v) methanol]. The culture was grown in a baffled flask covered with 3 layers of cheesecloth to facilitate aeration, and incubated in an orbital shaking incubator at 30ºC for 4 days to allow for protein expression. The culture was supplemented with 0.5% (v/v) methanol daily to induce protein expression. Cells were pelleted (6000×g, 10 min, RT) and supernatants were used for purification of protein.

3.2.3 Purification of recombinant *Tvi*CATL and *Tco*CATL

Three phase partitioning (TPP):

Three phase partitioning was carried out according to Section 2.2.3. However, the *P. pastoris* expression supernatants were filtered through Whatman No. 1 filter paper and adjusted to pH 4.2 with phosphoric acid prior to the addition of ammonium sulfate and t-butanol. pH adjustment allowed for CATL-like peptidases to be processed autocatalytically to mature proteins. Tertiary-butanol [30% (v/v) final concentration] and ammonium sulfate (40% of final volume) were added and stirred until completely dissolved. Supernatant mixtures were centrifuged (6000 x *g*, 10 min, 4ºC) in a pre-chilled spin-out rotor. The protein layers were collected at the interface between the t-butanol and aqueous phases and re-dissolved in dialysis buffer (PBS, comprising 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.76 mM KH_2PO_4, pH 7.2).

Cation exchange chromatography:

*Tco*CATL was further purified using cation exchange chromatography. A sulfopropyl (SP) Sephadex C-25 cation exchange column (15 x 100 mm) was equilibrated with 50 mM sodium acetate buffer, pH 4.2. Dialysed protein was cycled overnight through the column at 4ºC (0.5 cm/min). Following cycling, the column was washed with 5 column volumes of 50 mM sodium acetate buffer, pH 4.2. Bound protein was eluted with a linear NaCl gradient from 0 to 1 M in 50 mM sodium acetate buffer, pH 4.2 over 5 column volumes. Elution of protein was monitored by measuring the absorbance of each fraction at 280 nm. The column was washed with 1 M NaCl (5 column volumes) to remove traces of bound proteins. The column was regenerated using one column volume each of 0.1 M HCl and 0.1 M NaOH. Sodium acetate buffer (50 mM, pH 4.2, 5 column volumes) was further added to re-equilibrate the column.

Molecular exclusion chromatography:

*Tco*CATL was further purified by molecular exclusion chromatography using a Sephacryl S-200 HR column as described in Section 2.2.3.

Concentration of protein:

Protein was concentrated by ultrafiltration (Centriprep Ultrafiltration device with YM-10 membrane). Once protein samples were centrifuged (3000 x *g*, 30 min, 4ºC), the filtrate in the filtrate collector as well as the concentrated sample were collected and stored at -80ºC until analysed by SDS-PAGE.

Active site titration and inhibition of Tco*CATL*:

The determination of the concentration of active T*co*CATL was done by active site titration with E-64 as described in Section 2.2.7.

3.2.4 Expression of pyroglutamyl peptidase (PGP)

Glycerol stocks of recombinant *E. coli* BL21 (DE3) containing pGEX4T1-PGP (Mucache, 2011) were streaked onto 2x YT plates [1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl, 15 g/l bacteriological agar] containing ampicillin (100μg/ml) and incubated overnight at 37ºC. Single colonies were picked and grown in liquid 2x YT (100 ml) overnight at 37ºC. The overnight culture was transferred into fresh 2x YT liquid media (900 ml) containing ampicillin (50 μg/ml). The culture was incubated at 37ºC until an

OD_{600} of 1 was reached whereupon IPTG (0.1 M) was added to induce expression. The culture was incubated for a further 4 hours at 37ºC. Cells were pelleted by centrifugation (2000 x *g*, 10 min, 4ºC) followed by resuspension of cells in 20 ml PBS-Tween pH 7.2 [PBS containing 0.1% (v/v) Tween-20]. To facilitate lysing of the cells, lysozyme was added (1 mg/ml) and cells were freeze-thawed (-20ºC; RT) followed by sonication on ice for 4 x 30 s with 10 s intervals. Lysed cells were centrifuged (5000 x *g*, 10 min, 4ºC). The supernatant was filtered through Whatman No. 1 filter paper and the peptidase inhibitor, E-64 (1 mM) was added.

3.2.5 Purification of PGP

Swollen lyophilised glutathione S-transferase affinity resin (1 ml), packed in a Poly-Prep® chromatography column (Bio-Rad, USA), was equilibrated with 20 column volumes of PBS (pH 7.4). The filtered supernatant was cycled overnight (4ºC) through the column. To remove all unbound material, the resin was washed with 15 column volumes of PBS containing 1% (v/v) Triton X-100. Following washing, the column was equilibrated with 20 column volumes of thrombin cleavage buffer (20 mM Tris-HCl, pH 8.4, 150 mM NaCl, 2.5 mM $CaCl_2$). Thrombin (5 U) was mixed with thrombin cleavage buffer (1 ml) and added to the resin. The thrombin was incubated overnight in the column with gentle rocking (RT). The thrombin cleaved the PGP from the GST fusion tag and PGP was directly collected from the column. The column was washed with a further 10 column volumes of thrombin cleavage buffer to collect any residual PGP left in the column. Bound GST and any uncleaved fusion PGP were eluted from the resin with 10 mM reduced glutathione (10 ml in PBS, pH 7.4). Fractions containing PGP, as analysed by SDS-PAGE, were pooled and concentrated. The addition of 5 column volumes of 100 mM sodium borate buffer containing 500 mM NaCl, pH 8.0 followed by 5 column volumes of distilled water, 5 column volumes 100 mM sodium acetate buffer containing 500 mM NaCl, pH 4.0 and a further 5 column volumes of distilled water, was used to regenerate the column. The resin was stored in PBS with 0.2% (w/v) NaN_3 at 4ºC.

3.2.6 Enzyme activity of PGP

PGP (1.5 ng) diluted in Brij-35 [0.1% (v/v)] was incubated in enzyme assay buffer (100 mM Tris-HCl buffer, pH 8.0) at 37°C for 10 min. pGlu-AMC was added (0-100 μM) and fluorescence was measured ($Ex_{360\ nm}$ and $Em_{460\ nm}$) over 10 cycles in the Fluostar Optima

spectrophotometer (BMG Labtach, Germany). Enzyme activity was measured against blanks (0-100 μM substrate in enzyme assay buffer).

3.2.7 Quantification of protein

Purified *Tco*CATL, *Tvi*CATL and PGP were quantified by the methods described in Section 2.2.4. Protein was also quantified using a method that involved running equal volumes of BSA standards (0.125 – 2 mg/ml) and purified recombinant protein sample on a 12.5% reducing SDS-PAGE gel. Proteins bands were visualised by staining with Coomassie blue R-250 staining and destaining as described in Section 2.2.5 (Fig 3.1). The concentration of recombinant protein was estimated using the "density at tool counter" operation on the Quantity One software of the VersaDoc imaging system (Bio-Rad, USA).

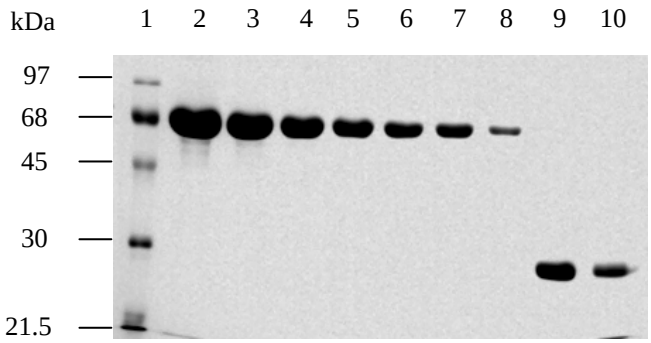


Figure 3.1 Direct visualisation on a reducing 12.5% SDS-PAGE gel for the estimation of recombinant protein concentration. Lane 1, molecular weight marker; lanes 2-8, BSA standards (2-0.125 mg/ml); lane 9, 1/10 dilution of purified PGP; lane 10, 1/20 dilution of purified PGP.

3.2.8 SDS-PAGE gel analysis of recombinant protein

The SDS-PAGE method (Laemmli, 1970), was used to visualise and analyse protein samples as described in Section 2.2.5.

3.2.9 Western blot analysis of recombinant protein

Western blots of recombinant proteins were carried as described in Section 2.2.6 using primary antibodies as anti-*Tco*CATL N-terminal IgY for *Tco*CATL and anti-ΔFLM *Tvi*CATL IgY for *Tvi*CATL. The secondary antibody used was rabbit anti-IgY-HRPO conjugate.

3.2.10 Optimisation of *Tco*CATL crystallisation conditions

Prior to designing a screen for *Tco*CATL crystallisation conditions, a precrystallisation trial (Rupp, 2009) was done. One µl of protein was mixed with 1 µl of the highly concentrated precipitant, PEG 6000 [30% (w/v)] on parafilm (Pechiney Plastic Packaging, IL, USA). The drop was observed immediately under an Olympus CKX41 Inverted Microscope (Wirsam, South Africa). Precipitation was an indicator that the protein was sufficiently concentrated to crystallise. Optimisation studies were initially carried out for *Tco*CATL. The Protein Data Bank (http://www.rcsb.org/pdb/home/home.do) was used to search for proteins with similar sequences to *Tco*CATL that had been crystallised. Proteins identified included native papain (Kamphuis et al., 1984) native papain in complex with a recombinant cysteine protease inhibitor derived from *T. b. brucei* (Alphey and Hunter, 2006), native papain-E-64 complex (Varughese et al., 1989), human CATL-E-64 complex (Fujishima et al., 1997), recombinant *Tcr*CATL in complex with Z-Phe-Arg-CH_2F (McGrath et al., 1995) or with K11777 (Kerr et al., 2009), recombinant *Tbr*CATL catalytic domain in complex with K11777 as well as falcipain-3 (cysteine peptidase from *Plasmodium falciparum*) in complex with K11017 (Kerr et al., 2009). The crystallisation conditions for each protein were used to design a crystallisation screen (Fig. 3.2.). CrysPred predictor (Kantardjieff and Rupp, 2004) was also used to explore possible pH values, based on the proteins' pI values to incorporate into the screen design (Fig. 3.2). Reagents made to be dispensed into a deep well block (Section 3.2.12) included sodium citrate (1.6 M), imidazole (1 M, one stock at pH 6.5 and the other at pH 8.5), ammonium sulfate (4 M), HEPES (1 M, one stock at pH 6.5 and the other at pH 8.5) and 50% (w/v) PEG 6000. A crystallisation plate (described in Section 3.2.11) was prepared for the screen. Two different ratios of protein were tested in the subwells of the crystallisation plate (Fig. 3.2).

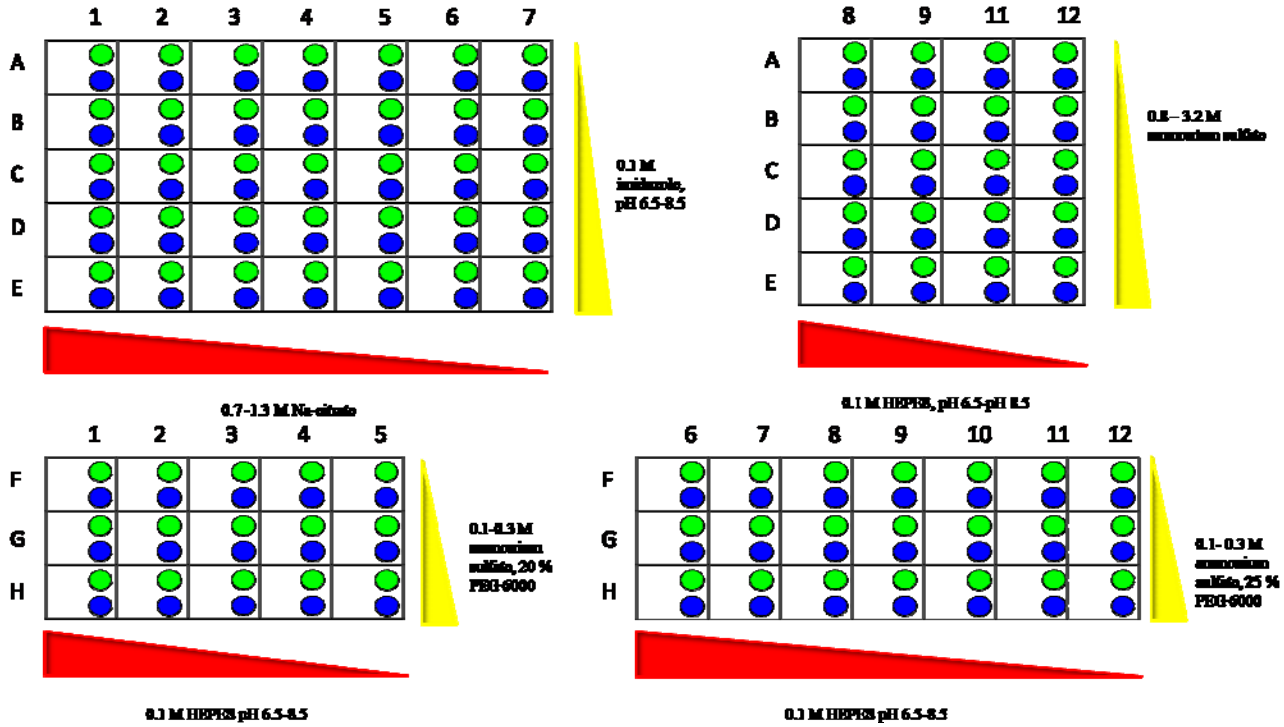


Figure 3.2. Screen design for *Tco*CATL in crystallisation plate. The top half of the plate (A-E) was used to screen *Tco*CATL with Na-citrate and ammonium sulfate in imidazole (increments of 0.5 in pH values) and HEPES (increments of 0.67 in pH values) buffers respectively. In the bottom half of the plate (F-H) different PEG 6000 concentrations were tested in HEPES pH values (column 1-5, increments of 0.5 in pH values; column 6-12, increments of 0.17 in pH values) and ammonium sulfate concentrations. Protein:crystallant ratios tested were 50:50 (green coloured subwell) and 60:40 (blue coloured subwell). The total volume of the crystallisation drop was 0.6 µl.

3.2.11 Setting up crystallisation screens for recombinant proteins

The Rigaku CrystalTrak™ application software allows capturing and recording of information on protein crystallisation experiments. In order to manage crystallisation screen designs, a crystallisation plate needs to be "created" in CrystalTrak™. Crystallisation plates for *Tco*CATL, *Tco*CATL mutants, *Tvi*CATL and PGP recombinant proteins were set up for screens described in Section 3.2.10 and later for the JCSG+ and PACT commercial screens (Molecular Dimensions, United Kingdom). Crystallisation plate information for screening the crystallisation conditions for each recombinant protein was entered into the *Plates* view in the *Setup* tab (Fig 3.3). This allows stored images in the CrystalTrak™ results section to be filed under a plate name. Plates were named according to the screen used and protein screened (e.g. JSCG+*Tco*CATL). The commercial screen designs used were selected from the list of commercial screen designs stored on the CrystalTrak™. For personalised optimisation screens, screens were designed by using the screen design option on the CrystalTrak™ (described in Section 3.2.12). A unique barcode

ID was assigned to the plate. The barcode was printed and stuck to the plate to allow for the correct storage of the plate in the Rigaku Gallery 160 Plate and scheduled imaging Hotel™. As there are two subwells for each of the 96 reservoirs in the sitting-drop plate, two different ratios of protein:precipitant were tested. The ratios of 67:33 and 50:50 for protein:precipitant were selected for the first and second subwell respectively. The total volume selected for the drops in each subwell was 0.6 µl.

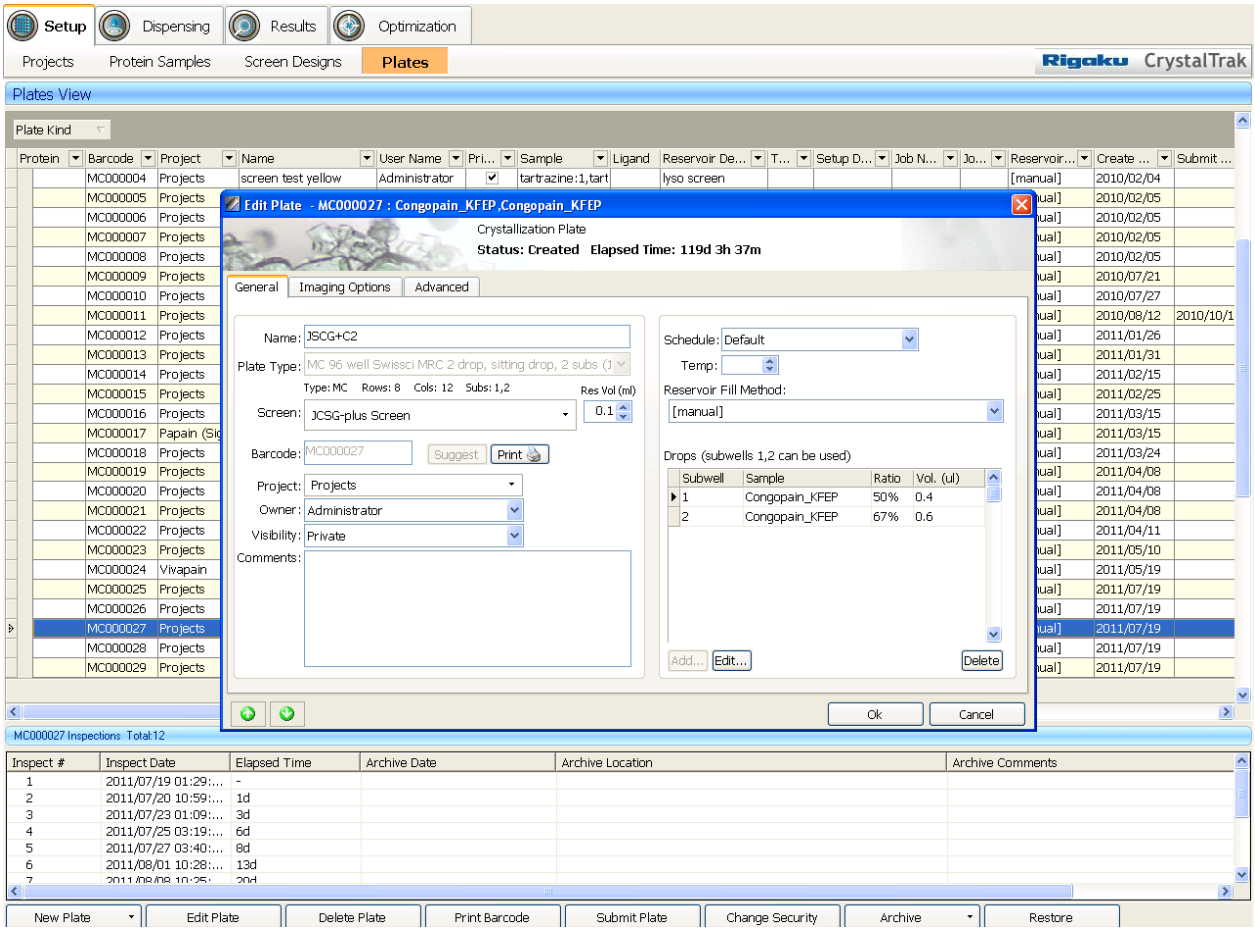


Figure 3.3 Creating and managing a crystallisation plate on CrystalTrak™. For each plate the name, screen design and barcode were created. The amount of protein and crystallant for each subwell of the plate were also noted.

3.2.12 Setting up deep well blocks containing crystallisation screen solutions on CrystalTrakTM for export to Alchemist II™

A deep well block (96 x 1 ml wells) was used to prepare the crystallisation optimisation screen. Commercially available screen designs are already stored on the CrystalTrak™ and do not have to be designed. For personalised optimisation screens, the CrystalTrak™ is used to design the composition of each well in the deep well block. For example for the optimisation of crystallisation conditions for *Tco*CATL (K39F; E44P), two designs, one

for PEG 3350 and the other for PEG 6000, were made. Design names were assigned as PEG6000_OPT and PEG3350_OPT. The 96 wells, characteristic of the deep well block, appeared on the right hand side of the screen (Fig. 3.4). Wells were selected and PEG and buffer pH gradients were "added". The screen design was exported to the Rigaku Alchemist II™ software. The Alchemist II™ system is a robotic liquid dispensing unit that is able to dispense liquid into a deep well block from stock solutions.

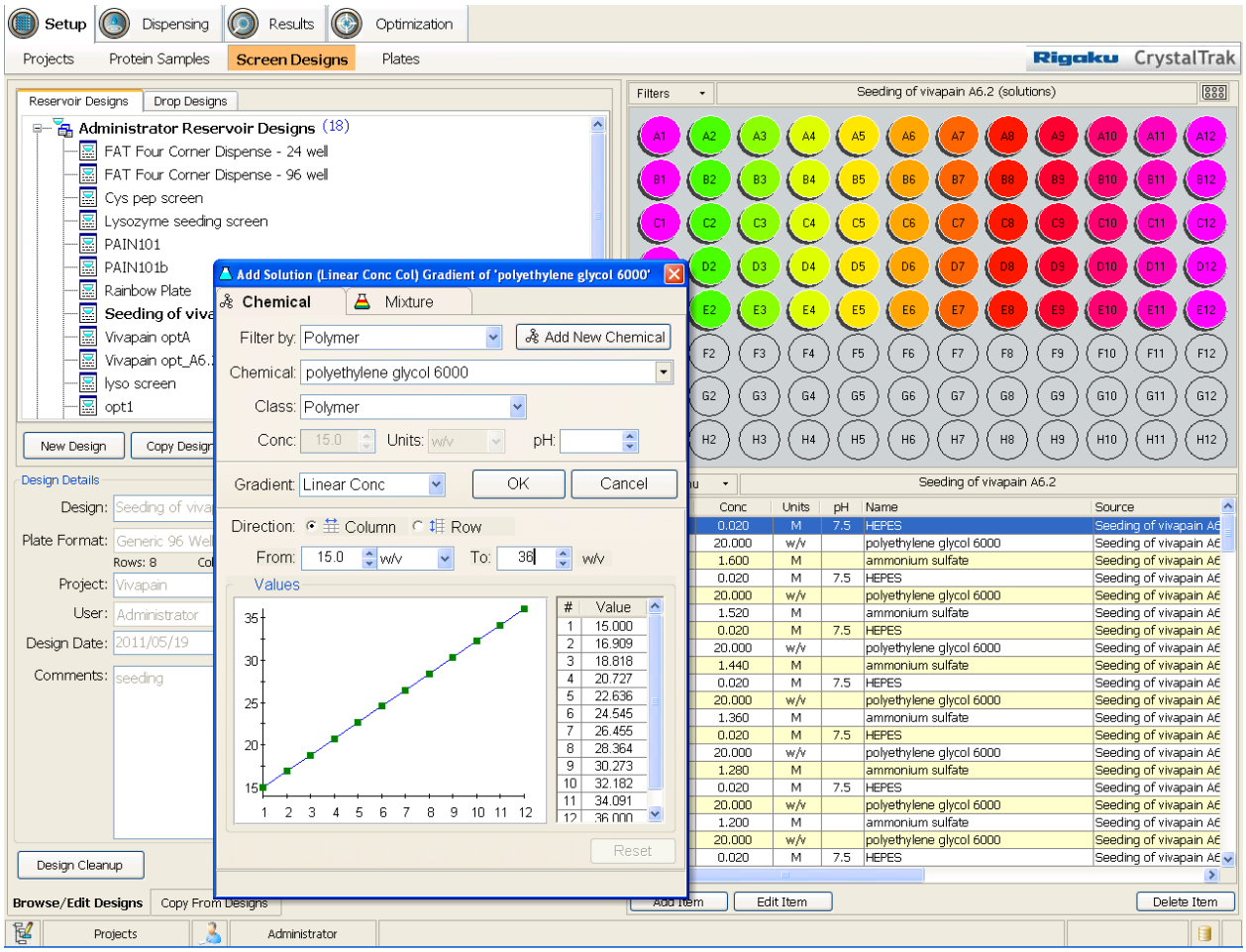


Figure 3.4 Screen design on the CrystalTrak™ software. The 96-wells, characteristic of the wells of the deep well block are highlighted and chemicals are added for gradient design.

Stock solutions were made, which included sodium citrate buffer (1 M, one at pH 5.0 and one at pH 5.5), 50% (w/v) PEG 6000 and 50% (w/v) PEG 3350 and dH_2O. The Alchemist II™ system was then used to dispense liquid into the deep well block. A crystallisation plate, with information about the plate stored in the CrystalTrak™ software and the Phoenix™ was used to dispense the solutions comprising the screen (from the deep well block wells) and the protein into the crystallisation plate. This plate, like the initial

commercial screening plates was loaded in the Gallery 160 Plate Hotel™ and scheduled inspections of crystallisation drops from plates were recorded.

3.2.13 Using the Phoenix™ for liquid dispensing into the crystallisation plate

The dispensing of the protein and crystallant screen solutions into the crystallisation plate was done by the multiposition, high-throughput liquid-handling robotic stage called the Phoenix™. The Phoenix™ system has Microsoft Windows-based software that allows for creating protocols from a list of commands. A specific protocol for the initial screening of the crystallisation conditions for the recombinant proteins using the designed screens was created (Fig 3.5). Once the protein and crystallant were dispensed into the 96-well sitting-drop plate, the plate was immediately sealed with sealant (UV sealant) and subsequently loaded into a cassette in the Rigaku Gallery 160 Plate Hotel™. The plate was lodged into a cassette within the Rigaku Gallery 160 Plate Hotel™. The crystallisation plate is scanned and the plates barcode is added to the inventory list and the schedule list on the DesktopMinstrel™ system.

3.2.14 Scoring images from screening recombinant proteins

The images of plates taken at scheduled time intervals are shown on the DesktopMinstrel™ UV system can be annotated using the CrystalTrak™ Imaging and scoring feature in the *Results* tab. Scoring is divided into four categories on the CrystalTrak™: clear, precipitate, crystal and other (Fig. 3.6). Based on precipitation results, crystallisation drops can either have “good” or “bad” precipitate. When the precipitation is visible under UV light, the drop can be scored as a good precipitate. If heavy precipitate is seen in the drop under visible light and nothing is observed under UV light, the drop generally contains a bad precipitate. Crystals can be scored to their size or the amount nucleation that has taken place. Although quite subjective, “Bad form” crystals are usually overnucleated and form as a shower of microcrystals. “Harvestable crystals” can sometimes be badly formed but generally are large singular, three-dimensional crystals. Transparent clusters, round droplets, spherolite, phase separation or oil-like films are scored from the sub-categories of “Other” (Fig 3.6).

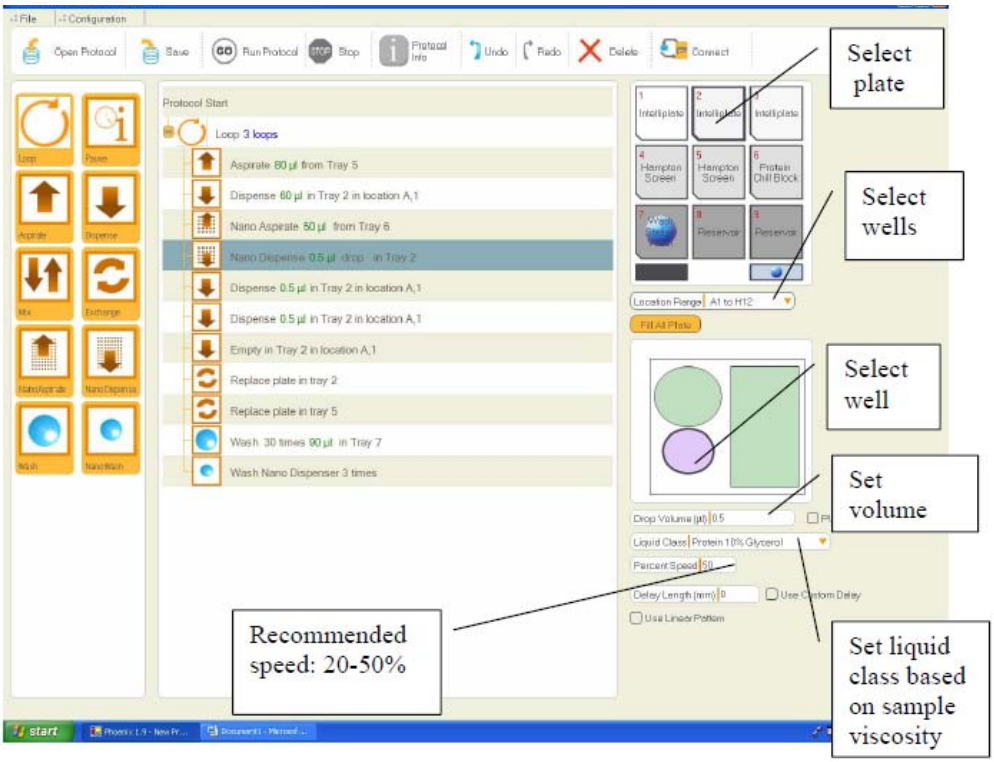


Figure 3.5 Creating a protocol using the Phoenix™ liquid handling software. A list of commands are situated on the left of the screen to choose from when making a personalised protocol. The right of the screen shows the layout of plates on the Phoenix™ as well as what the command will do with respect to each plate.

3.2.15 Setting up hanging drop crystallisation format for growth of *Tco*CATL (K39F; E44P) crystals

In order to transfer *Tco*CATL (K39F; E44P) crystals into capillary tubes for shipment to Professor Wolf-Dieter Schubert (University of Western Cape) for data collection, the crystals were grown using the hanging-drop method. A volume of 900 µl of crystallant [18.8% (w/v) PEG 6000, 0.13 M Na-Citrate, pH 5.5] was placed into a well of a 24-well cell culture plate. Two µl of each crystallant and *Tco*CATL (K39F; E44P) were placed on a microscope coverslip. The coverslip was inverted and placed on top of the well. To ensure an air-tight seal between the coverslip and the well, a small amount of silicon was placed around the rim of the well before the coverslip was placed over the well. Crystals grown from the hanging drop were transferred into a capillary tube. The capillary tube was sealed on either ends with beeswax.

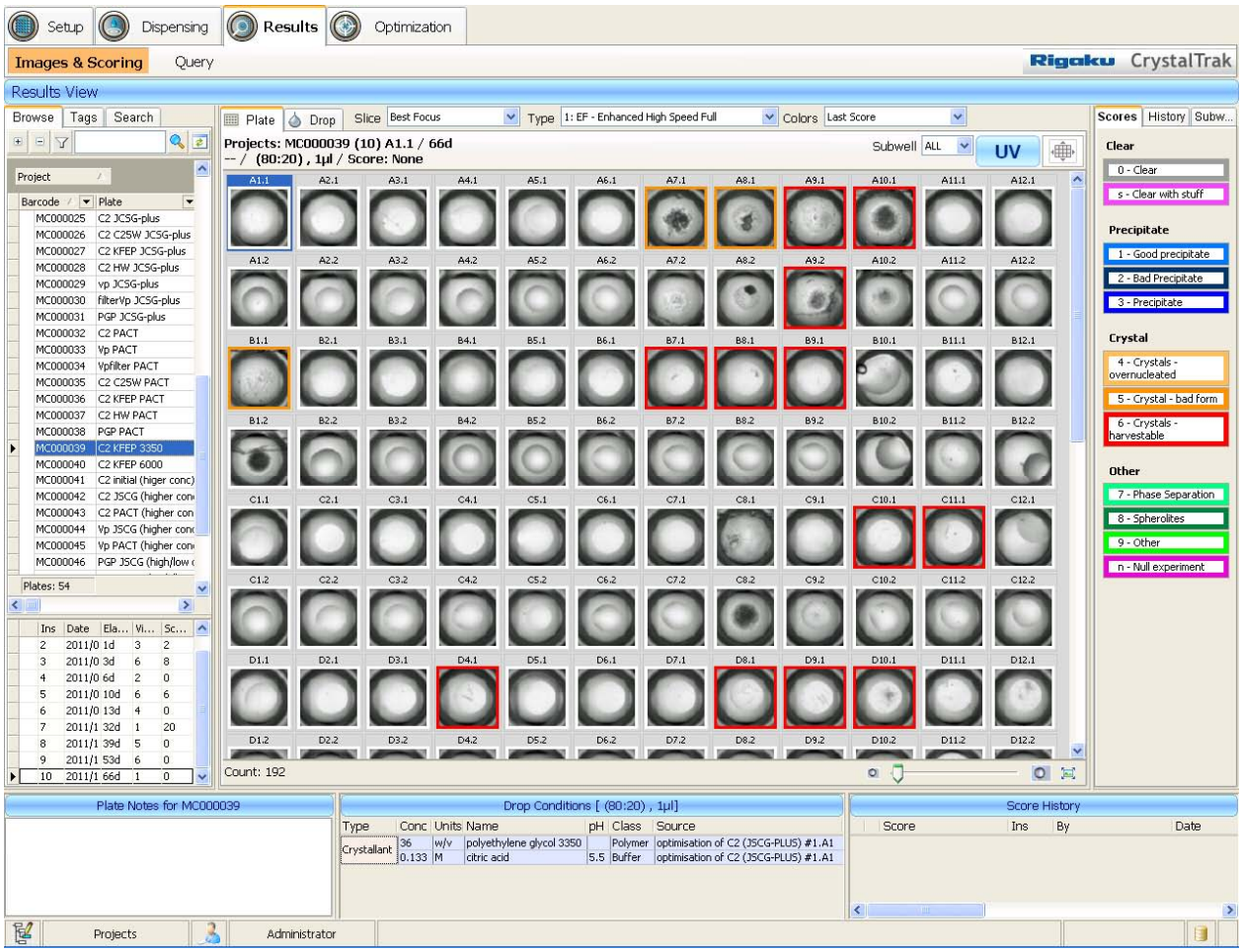


Figure 3.6 Scoring stored images on the CrystalTrak™. Images are scored according to the scores categorised as clear, precipitate, crystal or other, as shown on the left of the screen. Each of the scores are sub-divided into different characteristics of the score.

3.3 Results

3.3.1 Expression and purification of *Tco*CATL and *Tvi*CATL

*Tco*CATL was recombinantly expressed in *P. pastoris*. *Tco*CATL was observed as a band of approximately 27 kDa (Fig. 3.7, lanes 2, outlined by red box, and 3), showing that *Tco*CATL was actually processed to its mature form before being subjected to acidification to allow for autocatalytic cleavage of the propeptide (Boulangé et al., 2001; Serveau et al., 2003; Vernet et al., 1995). Three phase partitioning, as an initial purification and concentration step for *Tco*CATL, allowed for *Tco*CATL to optimally precipitate in 40% ammonium sulfate. At 40% ammonium sulfate precipitation, two other proteins of 36 and 38 kDa also precipitated (Fig. 3.7, lane 3). The presence of higher molecular weight contaminants required further purification steps.

*Tco*CATL started eluting at approximately 200 mM NaCl on a SP Sephadex C-25 column as shown by higher A_{280} readings of the fractions (Fig. 3.8, panel A). Fractions 18-28 were analysed on a SDS-PAGE gel (Fig. 3.9, panel B). Purification of *Tco*CATL was achieved using cation exchange chromatography which separated the *Tco*CATL from higher molecular weight contaminants observed after TPP (Fig. 3.8, panel B). *Tco*CATL could be observed as a doublet of 26 and 27 kDa on the SDS-PAGE gel. The faint lower molecular weight bands of 21.5 and 18 kDa were speculated to be degradation products of *Tco*CATL resulting from autocatalytic activity.

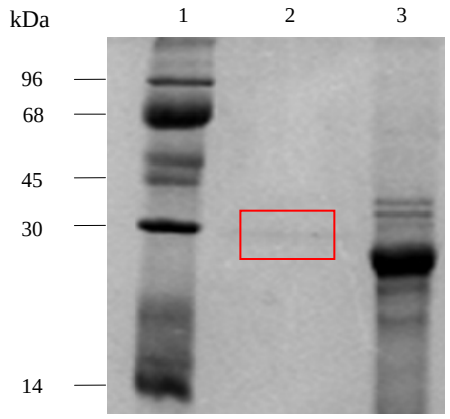


Figure 3.7 Purification and concentration of *Tco*CATL from expression supernatant by three phase partitioning (TPP). *Tco*CATL from *P. pastoris* expression supernatant was purified and concentrated by TPP, using 40% ammonium sulfate, and analysed on a 12.5% reducing SDS-PAGE gel. Lane 1, molecular weight marker; lane 2, *Tco*CATL from expression supernatant; lane 3, *Tco*CATL precipitated using TPP.

To further purify the *Tco*CATL sample, the fractions collected from cation exchange chromatography were pooled and concentrated by ultrafiltration before separation on the molecular exclusion chromatography column. Monitoring the absorbance of eluted samples at 280 nm and their hydrolysis of Z-Phe-Arg-AMC showed elution of *Tco*CATL at 27 kDa (Fig. 3.9). However, elution at a highly molecular weight has been previously shown. It was proven to be attributed to a dimeric *Tco*CATL conformation (Ndlovu, 2010). Lower molecular weight contamination was still present (Fig. 3.9), confirming that the lower molecular proteins could be degradation products of *Tco*CATL.

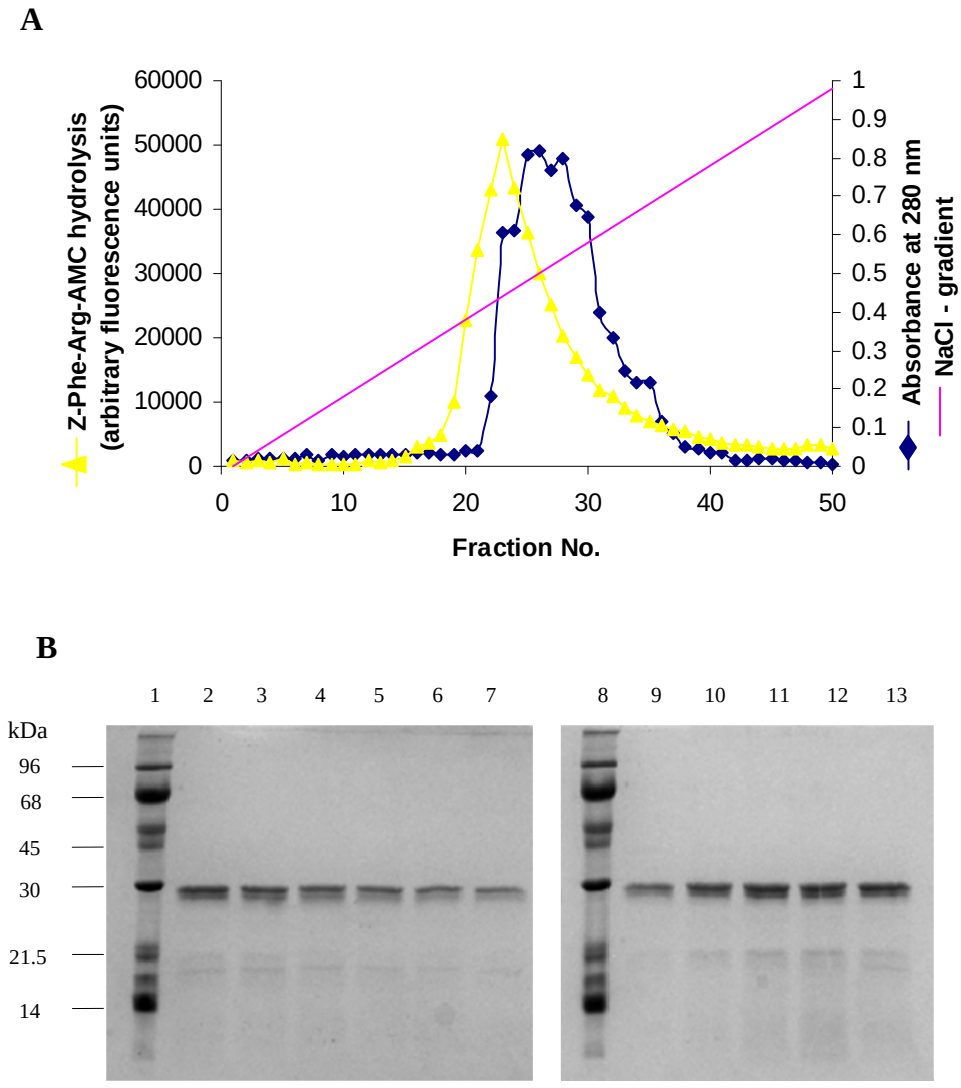


Figure 3.8 Purification of *Tco*CATL by cation exchange chromatography. Panel A, elution of *Tco*CATL from an SP Sephadex C-25 column (1.5 × 10 cm, 0.5 cm/min) using a 0-1 M NaCl (—) gradient in 50 mM sodium acetate buffer, pH 4.2. Fractions were monitored by reading the absorbance at 280 nm (♦) and measuring the hydrolysis of Z-Phe-Arg-AMC (▲). **Panel B**, analysis of the purity of the peak fractions by reducing SDS-PAGE (12.5%). Lanes 1 and 8, molecular weight marker and lanes 2-7 and 9-13, fractions 18-23 and 24-28 eluted from the column.

Fractions containing *Tco*CATL obtained from molecular exclusion chromatography were pooled and concentrated using ultrafiltration. The concentration of *Tco*CATL was calculated to be 7 mg/ml, giving a 7-fold higher yield than that obtained for the expression of *Tco*CATL in baculovirus (Boulangé et al., 2001). Both the ultrafiltration filtrate and the concentrated *Tco*CATL sample were analysed on a SDS-PAGE gel (Fig. 3.10, panel A). The 21.5 kDa band seen in the concentrated sample was recognised by the anti-*Tco*CATL N-terminal peptide IgY in a western blot (Fig. 3.10, panel B), verifying the autoproteolytic activity of *Tco*CATL under reducing conditions.

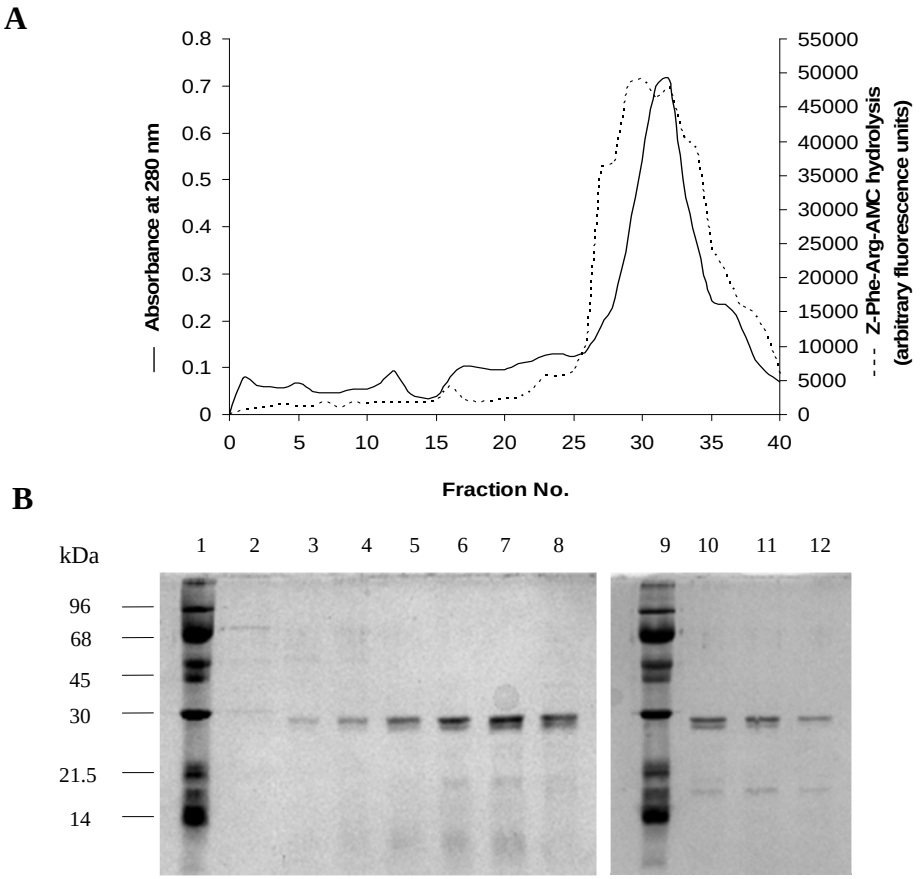


Figure 3.9 Purification of *Tco*CATL by molecular exclusion chromatography. Panel A, elution of *Tco*CATL from a Sephacryl S-200 column (25 x 650 mm, 12.2 cm/min) in 0.02 M Tris-HCl, pH 7.6, 0.02% (w/v) NaN_3. Fractions were monitored by reading the absorbance at 280 nm (—) and by measuring the hydrolysis of Z-Phe-Arg-AMC (---). **Panel B**, analysis of the purity of the peak fractions by reducing SDS-PAGE gel (12.5%). Lanes 1 and 9, molecular weight marker and lanes 2-8 and 10-12, fractions 26-32 and 33-25 from Sephacryl S-200 column.

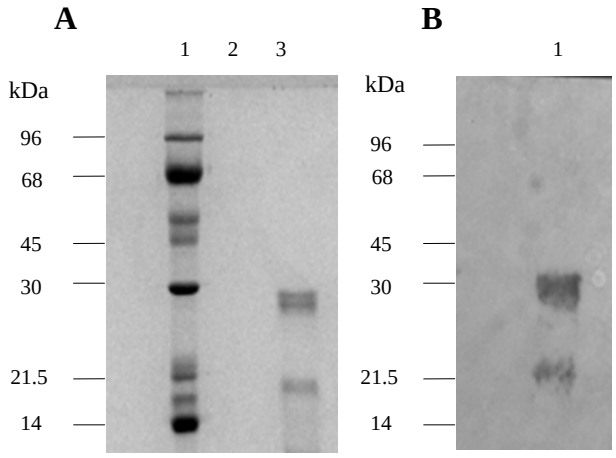


Figure 3.10 Analysis of concentrated *Tco*CATL by reducing SDS-PAGE gel and western blotting. Panel A, *Tco*CATL samples purified by molecular exclusion chromatography were pooled and concentrated by ultrafiltration and analysed on a reducing 12.5% SDS-PAGE gel. Lane 1, molecular weight marker; lane 2, ultrafiltration filtrate; lane 3, concentrated *Tco*CATL sample. **Panel B**, western blot of concentrated *Tco*CATL probed with anti-*Tco*CATL N-terminal IgY (10 μg/ml). Secondary antibody was rabbit anti-IgY (1:10 000 dilution). Development was achieved with H_2O_2 and 4-chloro-1-naphthol. Lane 1, concentrated *Tco*CATL sample.

Since co-crystallisation performed with inhibitors can be favourable for crystallisation and reduce protein microheterogeneity through autoproteolysis (Ostermeier et al., 1997), *Tco*CATL was incubated with a range of concentrations of E-64 (Fig. 3.11, panel A). This active-site titration allowed calculation of the concentration of active *Tco*CATL. Twenty % of the *Tco*CATL was active which amounted to a concentration 1.4 mg/ml active enzyme. Consequently 0.1 mM E-64 (molar excess of active *Tco*CATL) was added to the *Tco*CATL sample and analysed on a SDS-PAGE gel, resulting in a single band of 27 kDa (Fig 3.11, panel B), evidence of inhibition of the autocatalytic activity of *Tco*CATL.

*Tvi*CATL was expressed in the *P. pastoris* expression system. Analysis of the expression supernatant and the TPP sample showed that *Tvi*CATL expressed as two bands at approximately 29 and 33 kDa. As in the case of *Tco*CATL, autocatalytic cleavage of the propeptide occurred in the expression supernatant, prior to acidification for removal of propeptide (Fig. 3.12, panel A). Both bands were recognised when probed with anti-ΔFL*Tvi*CATL IgY (Fig. 3.12, panel B). Treatment with endoglycosidase H has previously shown that a single band is obtained following deglycosylation of the possible glycosylation site at Asn288 (Vather, 2010). No further purification after TPP was needed for *Tvi*CATL. The TPP sample was concentrated to 10 mg/ml, adequate for crystallisation studies.

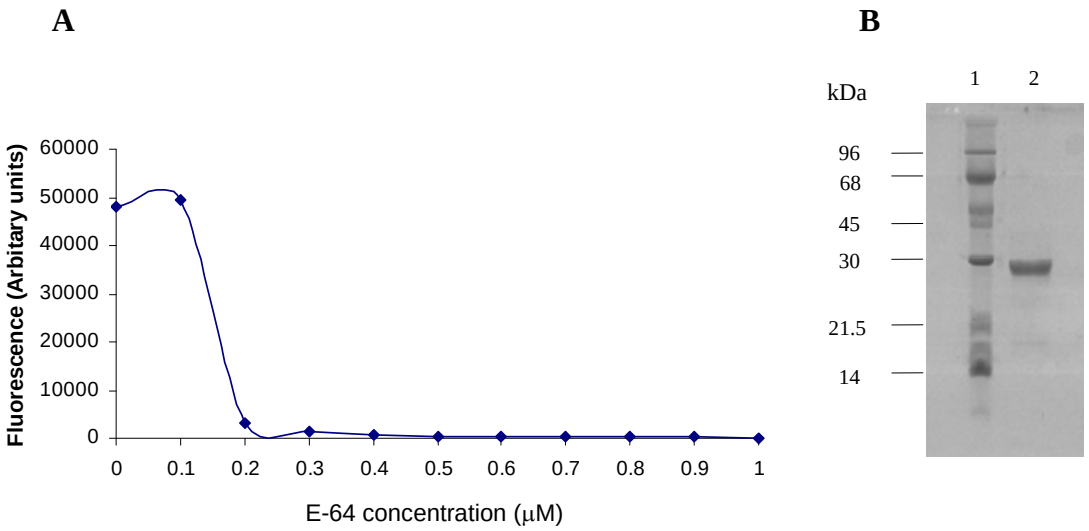


Figure 3.11 Active site titration of *Tco*CATL and reducing SDS-PAGE analysis of *Tco*CATL incubated with E-64. Panel A, *Tco*CATL (1 μM) was incubated with different concentrations of E-64 (0.1-1 μM) followed by measuring hydrolysis of Z-Phe-Arg-AMC. **Panel B**, analysis of *Tco*CATL inhibited with 0.1 mM E-64 on a 12.5% reducing SDS-PAGE gel. Lane 1, molecular weight marker; lane 2, inhibited *Tco*CATL.

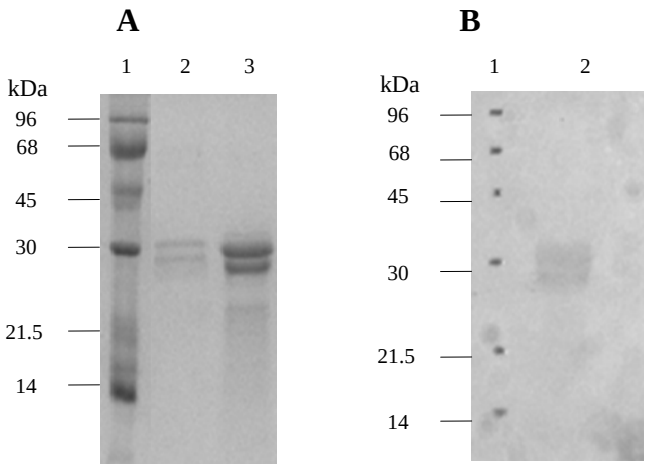


Figure 3.12 Purification and concentration of *Tvi*CATL from expression supernatant by three phase partitioning (TPP). Panel A, *Tvi*CATL from *P. pastoris* expression supernatant was purified and concentrated by TPP, using 40% ammonium sulfate, and analysed on a 12.5% reducing SDS-PAGE gel. Lane 1, molecular weight marker; lane 2, *Tvi*CATL from expression supernatant; lane 3, precipitated *Tvi*CATL after TPP. **Panel B**, western blot of concentrated *Tvi*CATL probed with anti-ΔFLM *Tvi*CATL IgY (10 μg/ml) followed by rabbit anti-IgY-HRPO conjugate (1:10 000 dilution). Development was achieved with H_2O_2 and 4-chloro-1-naphthol. Lane 1, molecular weight marker; lane 2, concentrated *Tvi*CATL sample.

3.3.2 Expression and purification of PGP

E. coli JM109 cells carrying the recombinant pGEX4T1-PGP were grown in 2x YT liquid media and induced with IPTG. PGP expressed as a 51 kDa protein as can be seen when comparing induced and uninduced expression samples (Fig. 3.13, lanes 2 and 3). PGP was released from GST by thrombin cleavage during the purification which resulted in a 28 kDa PGP sample (Fig. 3.13, lanes 5-10). Free GST was eluted at 25 kDa upon the addition of glutathione to the column (Fig. 3.13, lanes 19 and 20 and lanes 22-29). Samples containing PGP were reapplied to the regenerated GST-bind column as a means to remove contaminating proteins of sizes 35, 33, 25 and 20 kDa that coeluted with the first few purified PGP fractions (Fig. 3.13, lanes 5-10). PGP was concentrated by ultrafiltration and was analysed by direct visualisation on an SDS-PAGE gel with BSA standards (Fig 3.1). The purified PGP sample was calculated to be 7.5 mg/ml, adequate for crystallisation studies. The PGP purified sample was also active against the pGly-AMC substrate (results not shown).

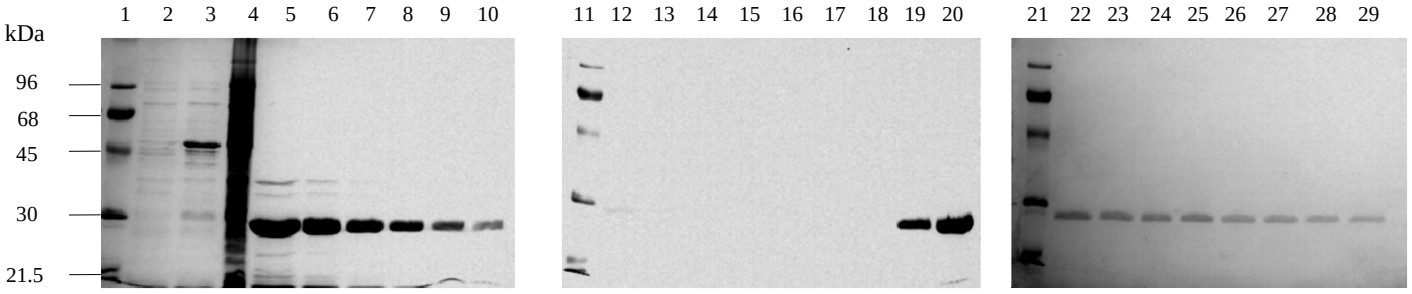


Figure 3.13 Analysis of PGP purification on a reducing 12.5% SDS-PAGE gel. Glutathione S-Transferase affinity chromatography purification. Lane 1, 11 and 22, molecular weight marker; lane 2, uninduced recombinant *E. coli* JM109 lysate; lane 3, induced recombinant *E. coli* JM109 (DE3) lysate; lane 4, unbound lysate; lanes 5-15 eluted PGP fractions; lanes 16-29, elution of GST.

3.3.3 Interpretation and scoring using a manual screen of *Tco*CATL crystallisation conditions

Concentrated and purified recombinant *Tco*CATL precipitated in the precrystallisation trial, proving to be concentrated enough for crystallisation screening. The visible and UV light images for each drop in the manual crystallisation screen were analysed every second day. The five major aspects that were compared were the concentration range of sodium citrate (0.7-1.3 M), the pH range of imidazole (pH 6.5-8.5), the concentration range of ammonium sulfate (0.8-3.2 M), the pH range of HEPES (pH 6.5-8.5) and the concentration of PEG 6000 at lower ammonium sulfate concentrations. After 20 days, images from crystallisation drops were either ‘Clear’ or had ‘Phase separation’ (Fig. 3.14). ‘Clear’ was allocated to crystallisation drops where no nucleation event had occurred and the visible and UV light images for crystallisation drop remained clear. ‘Phase separation’ was shown by small droplets within the crystallisation drop. The concentration of sodium citrate affected the level of phase separation more than the pH range of imidazole. Phase separation was seen across the pH range of imidazole. However the level of phase separation steadily decreased over the increasing concentration of sodium citrate. Phase separation was found across the pH range of HEPES (pH 6.5-8.5) and ammonium sulfate (0.8-3.2 M). The most promising condition was the phase separation found for the 60:40 ratio of protein: crystallant in the crystallising reagent of 1.4 M ammonium sulphate with 0.1 M HEPES buffer, pH 7.17 (Fig. 3.14, C and D). All crystallisation drops that contained PEG 6000 at 20 and 25% (w/v) over the HEPES pH range (6.5-8.5) and ammonium sulfate concentration range (0.1-0.3 M) were scored as clear. Whether the PEG 6000 or lower concentration levels of ammonium sulfate caused the crystallisation drop to remain clear is still to be determined.

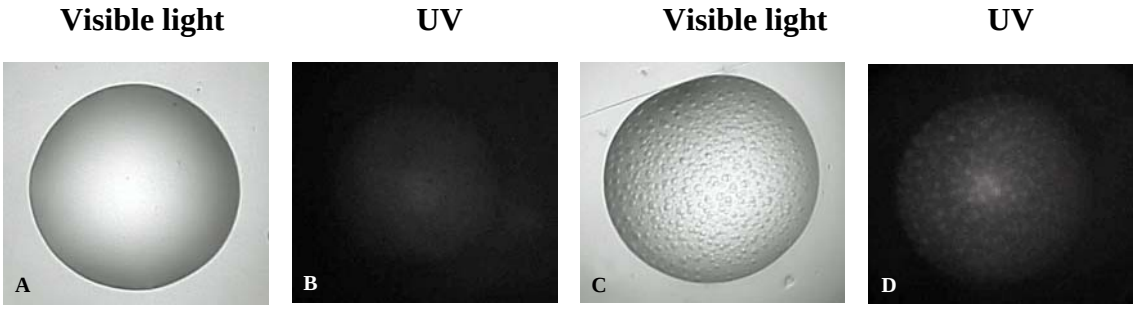


Figure 3.14 Scoring of crystallisation drop results from manual screen. Proteins were screened with conditions based on conditions of related proteins previously crystallised. After 20 days crystallisation drop images viewed under visible and ultraviolet (280 nm) light were scored according to appearance. A and B, clear drop; C and D, phase separation.

3.3.4 Scoring images from commercial crystallisation screens

The purified and concentrated *Tco*CATL, *Tco*CATL mutants, *Tvi*CATL and PGP were screened using the JSCG+ and PACT commercial screens. The images were looked at weekly and on day 30, images were scored. Images were scored according to the scoring table available on the Rigaku CrystalTrak™ software. As a reference, images were also compared to published images (Zeelen, 2009). Similarly to the manual screening, if no nucleation event had occurred and the visible and UV light images for crystallisation drop remained clear, the 'clear' score was then allocated (Fig. 3.15, panels A and B). However a 'Good precipitate' was one where precipitate was seen in the crystallisation drop under visible light and the distinct protein precipitation absorbed ultraviolet (UV) light at 280 nm (Fig. 3. 15, panels C and D). On the other hand a 'Bad precipitate' was a dark protein precipitate observed under visible light which was not detected at 280 nm (Fig 3.15, panels E and F).

The two hits, using the PACT screen, from *Tco*CATL (K39F; E44P) that were found at 30 days after screening the six recombinant proteins with different screens, were scored as 'Crystals-bad form' (Fig. 3.15, panels G and H and I and J) due to their needle-like, one-dimension form. The crystallant from the two hits both contained 0.1 M sodium citrate buffer with a protein:precipitant ratio of 67:33. However, the one hit contained 20% (w/v) PEG 3350, with the sodium citrate buffer at pH 5.5 and the other contained 20% (w/v) PEG 6000, with the sodium citrate buffer at pH 5.0. The hits were generated at a temperature of *ca* 21°C. These were the only two conditions in the screen that had a pH of 5 and 5.5.

3.3.5 Optimisation of *Tco*CATL (K39F; E44P) crystals

To improve the one-dimensional crystals of *Tco*CATL (K39F; E44P), the pH, protein concentration and PEG concentration (for both PEG 3350 and PEG 6000) were optimised. Crystals were grown within 3 days under many of the conditions (Fig. 3.16). Although most of the crystals were still needle-like there were variations within the size and diameter of the crystals. Overnucleation seemed to be a problem due to the shower of microcrystals in some of the drops. The optimisation screen containing PEG 6000 yielded slightly broader crystals than that of the PEG 3350 [Fig. 3.16 (g and i)]. One of the crystals produced here (Fig. 3.16 [g]) had more dimension than the one-dimensional needle-like crystals seen in Fig 3.15 (G and I).

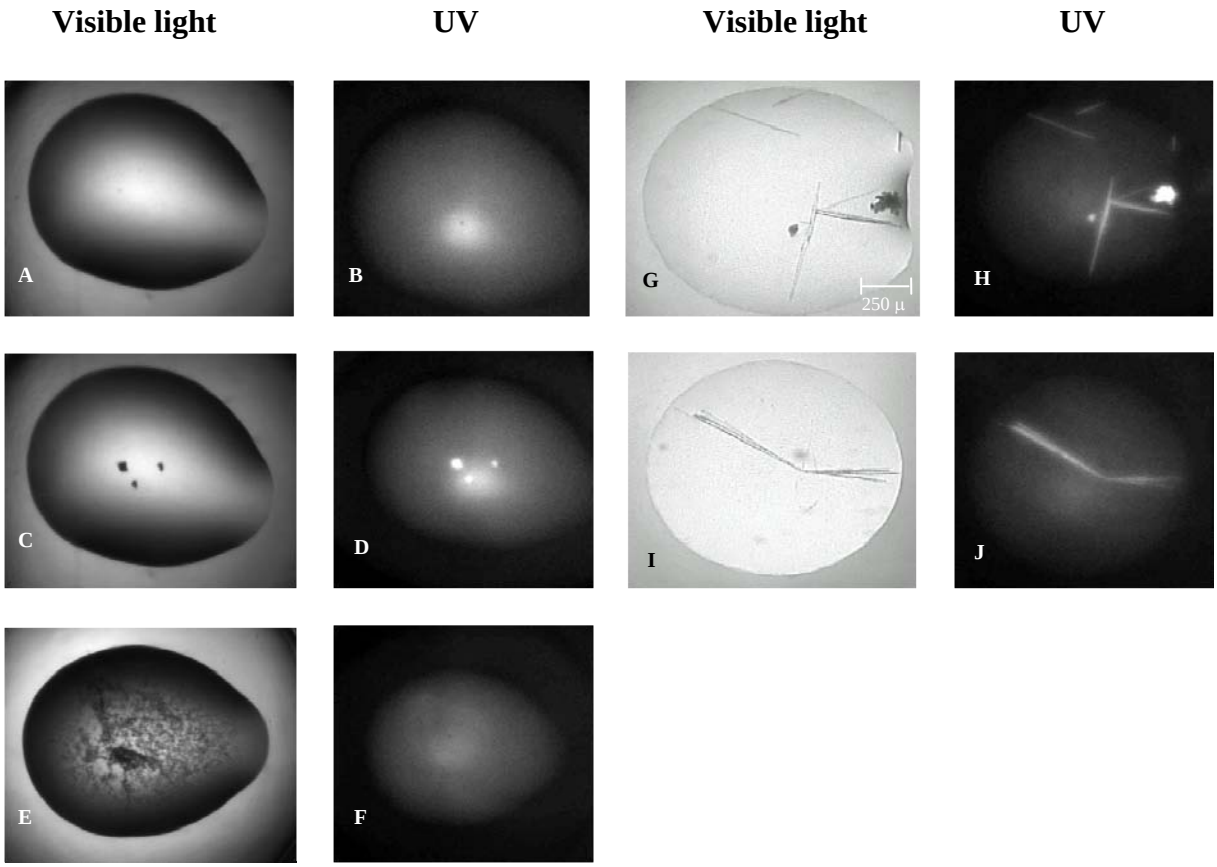


Figure 3.15 Scoring of crystallisation drop results from commercially available screens. Proteins were screened with different commercial screens and after 30 days crystallisation drop images observed under visible and ultraviolet (280 nm) light were scored according to appearance. A and B, clear drop; C and D, good precipitate; E and F, bad precipitate; G and H and I and J, crystals-bad form.

To check whether some of the crystals could diffract under the X-ray beam, they were sent to Professor Wolf-Dieter Schubert, University of the Western Cape. The sitting-drop

crystallisation plate was not ideal for sending the crystals in a safe and secure manner and the crystals needed to be transferred into capillary tubes (1 mm diameter). Crystals were grown using the hanging drop method for 3 days at 21°C, with protein:crystallant ratio of 80:20 and the total drop volume was 4 µl. The big size of the drop along with growing the crystals on the flat surface of the coverslip as opposed to the well of a sitting drop allowed for crystals to be transferred into capillary tubes via capillary action. Once the crystals were grown, an additional 3 µl of crystallant was added to the drop to bulk up the volume of the drop. The capillary tube was viewed under an Olympus CKX41 Inverted Microscope (Fig. 3.17). The capillary tube, sealed on either ends with beeswax, was sent to the University of Western Cape.

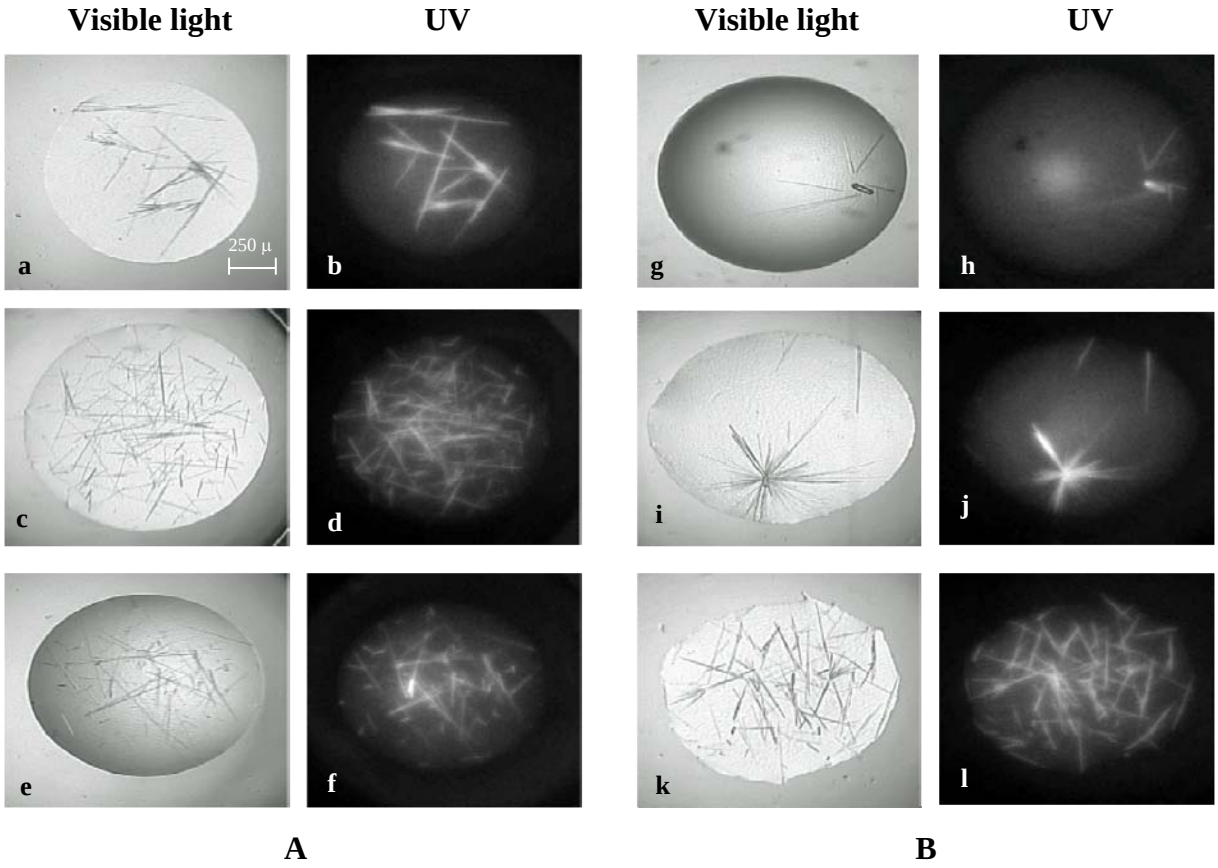


Figure 3.16 Crystals from optimisation crystal screen of *Tco*CATL (K39F; E44P). Images were taken of drops under visible as well as ultraviolet (280 nm) light, these are from day 3. All drops contained a ratio of protein:crystallant of 80:20 with the total drop volume being 1 µl and all crystals were grown at a constant temperature of *ca* 21°C. The concentration used for each condition was 10 mg/ml. Panel A, contained PEG 3350 and panel B contained PEG 6000. **Panel A**, a, b, 22% (w/v) PEG 3350, 0.1 M Na-citrate, pH 5.3; c, d, 19.8% (w/v) PEG 3350, 0.09 M Na-citrate, pH 5.1; e, f, 25% (w/v) PEG 3350, 0.13 M Na-citrate, pH 5.0. **Panel B**, g, h, 20% (w/v) PEG, 0.1 M Na-citrate, pH5.0; i, j, 18.8% (w/v) PEG 6000, 0.13 M Na-citrate, pH 5.5; k, l, 15% (w/v) PEG 6000, 0.13 M Na-citrate, pH 5.0.

A single crystal was tested for its ability to diffract. The diffraction was, however, weak. Reflections on the diffraction map (results not shown) were all near the origin indicating that it was a protein crystal. If it was salt, one would expect to see only a hand full of reflections distributed all over the detector. However, there were three reflections defining a right-angled triangle. This would appear to indicate that at least one of the angles of the unit cell equals 90° (Prof W.D Schubert, University of Western Cape, personal communication).

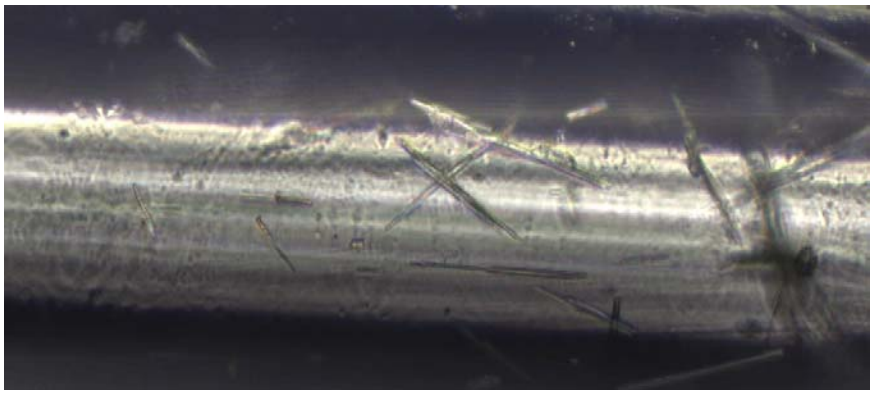


Figure 3.17 *Tco*CATL (K39F; E44P) crystals sealed in capillary tubes. Crystals were grown using the hanging drop technique at 21°C for 3 days. The protein:crystallant ratio was 80:20 in a total volume of 4 μl. Crystals were grown in 18.8% (w/v) PEG 6000, 0.13 M Na-citrate, pH 5.5. Crystals transferred into capillary tubes were viewed using the Olympus CKX41 Inverted Microscope (20x).

3.4 Discussion

Over most of sub-Saharan Africa, bovine trypanosomosis continues to be controlled primarily by trypanocidal compounds such as isometamidium chloride, homidium (bromide and chloride) and diaminazene (Holmes et al., 2004). The classification of trypanosomosis as being a neglected disease, by the World Health Organisation in 1975, has attracted researchers to the developments of more effective and new drugs. The vinyl sulfone inhibitor of *Tcr*CATL, K11777, was effective in curing or alleviating *T. cruzi* infection in preclinical proof-of-concept studies and has now entered formal preclinical drug development for Chagas' disease (McKerrow et al., 2009). This provides hope for chemical knockout of other pathogenic cysteine peptidases. Furthermore, the anti-disease vaccine approach targeting cysteine peptidase pathogenic factors (Authié, 1994) does, however, still hold promise. Based on the ability of trypanotolerant cattle to produce a higher antibody response to *Tco*CATL compared to trypanosusceptible cattle,

immunisation of cattle with baculovirus expressed *Tco*CATL provided evidence for cattle maintaining or gaining weight as well as less severe anaemia development (Authié et al., 2001). It is therefore imperative that the structure of pathogenic factors, such as the peptidases, be determined to understand the resultant structure has on function and design better vaccines and chemotherapeutics. In this part of the study, *Tco*CATL, *Tco*CATL mutants, *Tvi*CATL and PGP cysteine peptidases were recombinantly expressed and purified for crystallisation.

Both catalytic domains of *Tco*CATL and *Tvi*CATL as well as the *Tco*CATL mutants were expressed and purified to homogeneity at sufficiently high concentrations for screening of crystallisation conditions. Although full-length *Tco*CATL has been expressed at concentrations of 2-4 mg/l in yeast, the peptidase has shown to be highly unstable (Boulangé et al., 2011). It is not clear why the full-length peptidase is unstable, however, it is known that the unique C-terminal extension of trypanosomal cysteine peptidases does not play a role in stability or activity of the active enzyme (Cazzulo et al., 1997). Moreover, the aim of the expression of the full-length *Tcr*CATL by Alvarez et al (2002) was to determine the 3-D structure by X-ray diffraction of the C-terminal domain and to date, crystallisation has not been achieved possibly due to the instability of the C-terminus in the formation of ordered and packed crystals. The proregion was included in the constructs for these recombinant proteins as it plays a pivotal role in the correct folding of the expressed protein, as in the case for cathepsins (Vernet et al., 1995). Although failure to express *Tco*CATL in a bacterial system has been previously reported (Boulangé et al., 2001), *Tcr*CATL has been successfully expressed in a bacterial expression system as an active enzyme both in full-length and truncated forms (Eakin et al., 1993). Prior to *Tco*CATL expression in *P. pastoris* (Boulangé et al., 2011), *Tco*CATL was expressed in a baculovirus system, with a yield of 2 mg/ml (Boulangé et al., 2001). It was the relative complexity, incorrect folding of the expressed enzyme and cost of the baculovirus expression system that led to the use of the alternative *P. pastoris* expression system. Like the baculovirus system, and unlike the bacterial system, *P. pastoris* allows for post-translational modifications such as proteolytic processing, folding, disulfide bond-formation and glycosylation of protein (Wu and Letchworth, 2004).

Both *Tco*CATL and *Tvi*CATL could be observed as doublets on reducing SDS-PAGE gels. The size difference between the protein bands for *Tvi*CATL is more prominent than those

of *Tco*CATL. Previous deglycosylation tests with endoglycosidase H have shown that the double band of *Tvi*CATL is changed to a single band on a reducing SDS-PAGE gel, suggesting that Asn^{288}, a possible glycosylation site, becomes deglycosylated (Vather, 2010). Glycosylation sites for *Tco*CATL are proposed to be located in the proregion. The proregion contains four asparagine residues (Asn67, Asn78, Asn102 and Asn119) that could be potentially glycosylated. However, because most of the proregion is removed during maturation, glycosylation is a less likely reason for the formation of a double band. It was previously shown that the *Tco*CATL proenzyme mutant (C25A) is glycosylated by performing the endoglycosidase H test (Ndlovu, 2010). The less prominent second band seen in the purification process of *Tco*CATL could possibly be a result of the several cleavage sites used during the *in vitro* maturation process of the proenzyme, as described for other cysteine peptidases (Vernet et al., 1995). Similar findings have been illustrated for *Tcr*CATL (Eakin et al., 1993), *Tbr*CATL (papain-like cysteine peptidase from *T. b. rhodesiense*) (Caffrey et al., 2001) and *Tco*CATL 1, a closely related cysteine peptidase of *Tco*CATL found in the *T. congolense* genome that shares 90% sequence identity with congopain (Boulangé et al., 2001). The higher molecular weight protein (29 kDa) of the *Tco*CATL 1 doublet contained 3 amino acids upstream of the predicted maturation site, while the lower band (28 kDa) corresponded to the catalytic domain and 1 amino acid of the proregion (Boulangé et al., 2001). The size difference between the two protein bands is similar for *Tco*CATL in the present study. This likely demonstrates processing of the *Tco*CATL proenzyme along a cleavage sites in the proenzyme. In the case of *Tcr*CATL and *Tbr*CATL, N-terminal sequencing revealed that the band of higher molecular weight band contained also 3 amino acids upstream of the cleavage of the proenzyme. However, after complete inhibition of *Tco*CATL by E-64, a single band was observed on the reducing SDS-PAGE gel.

Similarly to *Tco*CATL, the *Tco*CATL mutants: *Tco*CATL (H43W) and *Tco*CATL (K39F; E44P), migrated as monomers of 27 kDa suggesting that their propeptides were cleaved off under acidic pH (Ndlovu, 2010). The inactive mutant *Tco*CATL (C25A), where the active site Cys residue was mutated does not undergo spontaneous cleavage when subjected to acidic pH (Boulangé et al., 2011). Full length recombinant *Tbr*CATL and *Tbr*CATL lacking the C-terminal extension were produced as already cleaved products in the yeast expression supernatant (Caffrey et al., 2001). Auto-activation had taken place in the expression medium and only a ‘clean up’ step was required at acidic pH to remove the very

small residual amounts of proenzyme. Analysis of the expression supernatant revealed that processing of the proregions of both *Tco*CATL and *Tvi*CATL did take place in the expression medium. As was the case for *Tbr*CATL, the manual change of pH to an acidic pH facilitated a 'clean up' of the *Tco*CATL and *Tvi*CATL proform. The secretion of mature *Tbr*CATL, *Tco*CATL and *Tvi*CATL contrasts with the production of CATF (Wang et al., 1998) and CATV (Brömme et al., 1999), which are produced as proenzymes in the same expression system. The *Tco*CATL mutants were also shown to be processed in the expression medium (Ndlovu, 2010). Maturation of *Tco*CATL has also been shown to occur at pH 8.0 in the presence of dextran sulfate, which for a papain-like enzyme is quite extraordinary (Serveau et al., 2003). It can merely be suggesed that *Tco*CATL could potentially be processed in the mammalian bloodstream of the infected host to its active form and as a result contribute to the pathogenesis of the disease (Serveau et al., 2003). In addition, *Tco*CATL, unlike *Tco*CATL 1 has shown to hydrolyse substrates over a wide pH range, from 4.0 to 8.5 (with the optimum being pH 6.0) (Lalmanach et al., 2002), adding further evidence of activity at physiological as well as lysosomal pH.

The homogeneity of each of the peptidases was evaluated after the purification process. TPP proved very successful in concentrating protein from the large volume expression supernatants. TPP is very reproducible, cheap and saves a large amount of time for the purification and concentrating process (Pike and Dennison, 1989). In addition the t-butanol allows the protein to float on the aqueous layer because t-butanol bound protein lowers the protein density (Pike and Dennison, 1989), which is advantageous in relation to general salting out. Both cation exchange and molecular size exclusion chromatography proved successful in removing contaminating proteins as evident by the SDS-PAGE gels. It has been mentioned in numerous studies that the purer the biological macromolecule the better the chance of obtaining crystals for diffraction studies (Benvenuti and Mangani, 2007; Derewenda, 2004; Gilliland and Ladner, 1996). For commercial preparations of lysozyme, three classes of impurities have been classified that influence crystallisation behaviour: contaminating proteins, small molecules and heterogeneous forms of the enzyme (Ewing et al., 1996). Heterogeneous forms of lysozyme appeared to have the greatest perturbing influence on crystallisation behaviour. Glycosylation has been said to interfere with crystallisation as the flexible and heterogeneous carbohydrate moieties can account for a significant fraction of the surface area of the protein, interfering with the crystallisation process (Derewenda, 2010). However, in a number of cases glycoproteins have been

successfully crystallised and the carbohydrate groups have actually facilitated the crystallisation process as the deglycosylated counterparts fail to crystallise (Aleshin et al., 1994; Mark et al., 2003). For the production of diffracting procathepsin L crystals, a T110A mutation was done to remove the O-linked glycosylation site (Coulombe et al., 1996). In the case of *Tbr*CATL crystals, the S172A mutation was made to remove the O-linked glycosylation site (Caffrey et al., 2001). As previously mentioned, *Tvi*CATL and *Tco*CATL (C25A) are both glycosylated and as a starting point in the study, an investigation was undertaken to see whether the glycosylated proteins were also able to crystallise.

The fact that *Tco*CATL, *Tco*CATL mutants and *Tvi*CATL were efficiently purified and yielded relatively high concentrations (> 5 mg/ml) allowed for crystallisation trials to be performed. The Rigaku crystallisation robotics has shown to increase the success rate of protein crystallisation, the major bottleneck to determining protein structure (Stevens, 2000; Buba et al., 2008). Prior to initial screening, the E-64 treated *Tco*CATL was to be present as a single protein band on an SDS-PAGE gel. There were no contaminating proteins or degradation products of *Tco*CATL present in the *Tco*CATL-E-64 mixture. *Tvi*CATL was characterised as being glycosylated and active (activity not shown). The *Tco*CATL mutants and PGP were not inhibited since they were present as a single band of protein. These are important biophysical characterisations that were recorded as future crystallisation studies could start with the manipulation of the biophysical character of the enzyme. However, other biophysical characterisations are available which concentrate more on protein folding and structure to determine initial crystallising ability. Two of the customarily employed methods for biophysical characterisation include emission fluorescence and light scattering (Derewenda, 2010). Emission light scattering allows measurement of changes in the fluorescence as a protein unfolds and/or undergoes important conformational changes induced by ligand/inhibitor binding. Dynamic light scattering is used in the monitoring of crystallisation trials because it detects the formation of aggregates or nuclei before they become visible using microscopy (Bergfors, 1999; Tessier and Lenhoff, 2003).

The phase separation produced by the different crystallisation conditions employed for the manual screening did provide a starting point of further screening. Phase separation can give rise to very high protein concentrations, with the protein concentrating in one of the

phases than the other, leading to extremely high protein concentrations and subsequently high states of supersaturation (McPherson, 2009). Furthermore, phase separation creates a boundary between the two solvents, serving as a nucleation center (McPherson, 2009). Besides the further manual screening that did not provide substantially different results to the original manual screening, commercial screening was done. Commercial screening provides a considerably large variation of crystallant reagents. JCSG+ and the PACT screens are known as the *evolution* and *intelligent design* respectively (Newman, 2009). The JCSG+, originally designed as the Core screen (Page et al., 2003), composed of 67 conditions that were extracted from 480 conditions, from 10 commercially available screens and tested on 539 proteins from *Thermatogamaritima*. This screen was expanded to the present 96 conditions (Page and Stevens, 2004). The PACT screen tests the effect of pH in 1 pH unit increments with multicomponent buffers. Furthermore, six cations and 12 anions are tested. The sitting drop vapour diffusion method employed for the crystallisation trials with commercial screens, using the Rigaku robotics, allowed for each of the 96 conditions from each of the two screens to be explored at two different concentrations of protein (two subwells for each of the 96 conditions on the sitting-drop plate). Effectively, this amounted to 384 conditions that were initially tested for each protein. The sitting-drop technique has benefits over the hanging-drop plating method such as cost and efficiency (Bolanos-Garcia and Chayen, 2009). The nanodrop needle of the Rigaku Phoenix^{TM} liquid dispensing unit, which works at a high throughput speed, is able to efficiently dispense protein on to the sitting-drop crystallisation plate using minimal protein sample, typically less than 5 μl.

The two initial hits that were found for the double mutated *Tco*CATL (K39F; E44P) were one-dimensional needle-like crystals. It has been said that needles are sometimes big enough to be used at the microfocusing beam at the synchrotron for X-ray diffraction (Zeelen, 2009). Optimisation results showed some form of overnucleation where many of the conditions produced a shower of smaller, microcrystals. However some crystals did have more dimension than others. The hanging-drop technique was used to grow crystals for the transfer into capillary tubes that could be sealed and sent off to the University of Western Cape for X-ray diffraction. The hanging-drop technique reduced the occurrence of hardware crystal adherence, as has been described previously (Bolanos-Garcia and Chayen, 2009). The weak diffraction pattern could have been attributed to the treatment of

the crystals, such as temperature and pressure variations as well as handling (Prof W.D Schubert, University of Western Cape, personal communication).

The number, ultimate size and possibly the quality of crystals depend on the mechanisms and the rates of nuclei formation (McPherson, 2009). Limited overnucleation by the adjustment of the physical and chemical properties of the crystallisation solutions used in screens are being pursued at present. Higher levels of supersaturation are required for nuclei to form than for subsequent crystal growth (Luft and DeTitta, 1999). The “labile zone” (from a crystallisation phase diagram) is the phase where nucleation takes place. After spontaneous nucleation, a lower level of supersaturation, referred to as the “metastable zone” is sufficient to support crystal growth. The ideal experiment should uncouple nucleation from crystal growth to satisfy the requirements of the two-phase events. It has been shown that seeding, through either streak seeding (Stura and Wilson, 1991) or microseeding (Luft and DeTitta, 1999) is the best method for uncoupling nucleation from crystal growth. Seeding should be carried out in the metastable zone of the phase diagram, which is generally achieved by placing seeds in a crystallisation drop of lower protein or crystallant concentration. Other chemical modifications that could also be evaluated involve additives such as detergents, reducing agents and heavy metal compounds. Mutations of residues involved in glycosylation and enzyme activity are being pursued. Furthermore, other inhibitors such as chloromethylketones and diazomethylketones for the cysteine peptidases are being used in complex with the enzyme for crystallisation screening. However every decision depends on the outcome of the previous crystallisation experiment which requires time to equilibrate and give positive and negative results. Moreover, between every experiment the protein sample ages which can have detrimental effects on the outcome of crystallisation.

Trypanosomal cysteine peptidases were recombinantly expressed and purified to homogeneity. The concentration of the purified proteins was suitable for crystallisation screening. Manual crystallisation screening of *Tco*CATL, based on conditions of related proteins previously crystallised, produced promising results. Some crystallisation drops showed phase separation which can give rise to higher protein concentrations and creates a boundary between the two solvents serving as a nucleation center. Furthermore, one-dimensional *Tco*CATL (K39F; E44P) crystals were grown in some crystallising reagents

of the PACT commercially available screen. Crystals collected after optimisation screening produced a weak diffraction pattern. Nonetheless, manipulations of crystallising conditions are still underway. The crystal structure of *Tco*CATL will be of major significance to the design of specific chemotherapeutic agents. In addition, understanding the dimeric conformation of *Tco*CATL is important for vaccine design as immune response are likely to recognise the dimer specific epitopes.

CHAPTER 4
GENERAL DISCUSSION

African animal trypanosomosis (Nagana) is a major limitation to livestock farming in Africa (Kristjanson et al., 1999). Ten million km^2 are infested with the disease vector, tsetse flies (Kabayo, 2002) where at least 46 million cattle and 70 million small ruminants are exposed to the risk of contracting trypanosomosis (Kristjanson et al., 1999). This widespread disease affects 40 African countries and is costing the African continent US $4.75 billion per year (Van den Bossche et al., 2010). In the acute stages of the disease, livestock usually develop fever and anaemia. The chronic stage occurs when cattle develop cachexia and both milk production and the oestrous cycle are impaired. The chronic phase can lead to death or in some cases self-cure or a persistent carrier state (Taylor and Authié, 2004).

Isometamidium chloride, homidium (bromide and chloride) and diaminazene are used, over most of sub-Saharan Africa, as the primary control for bovine trypanosomosis (Holmes et al., 2004). However, these drugs have been used for more than 50 years (Kroubi et al., 2011) and there is a need to develop drugs that are more therapeutic, combat resistance and possess fewer side effects. Other strategies for controlling the disease include the eradication or control of the tsetses fly populations. Besides targets and traps being used extensively in small agriculture settings (Vale, 1993), insecticides, such as pyrethroids, are also being used (Allsopp and Hursey, 2004; Eisler et al., 2003). Furthermore, other approaches to tsetse control have involved the sterile insect technique, a genetic population suppression approach that involves sustained, systematic releases of irradiated sterile male insects among the wild population (Aksoy, 2003).

Trypanosomes are able to continuously evade the host antibody response, with the generation of an inexhaustible collection of variable surface glycoproteins (VSGs) that cover the surface of the trypanosomes. Vaccination studies targeting VSGs of the parasite were proven to be ineffective as trypanosomes evade host antibody responses through the continuing changes in VSGs (Morrison et al., 1982). Other vaccination strategies have evaluated non-variable antigen-based vaccines. Trypanosomes express less variable

surface antigens such as invariant surface glycoproteins (ISGs). Although initial work on mice challenged with *T. b. brucei* after immunisation with ISG75 proved unsuccessful (Tran et al., 2008), other ISGs are being assessed. Alternatively, transmission blocking vaccines, aimed at reducing transmission of the trypanosome through immunisation of the host against insect parasite stages have been researched. However, the approach to this vaccine does have an ethical problem in that it does not protect the vaccinated individual against the infection but would however ultimately result in reduced numbers of infectious vectors. Anti-disease vaccines, have been considered as an promising approach in livestock (Authié, 1994). The purpose of an anti-disease vaccine is to reduce the symptoms of the disease rather than killing the parasite itself. Trypanotolerant cattle are able to protect themselves from the disease-associated symptoms (Authié et al., 2001). Trypanotolerant cattle mount an antibody response to the cysteine peptidase, *Tco*CATL. The artificial induction of antibodies against the peptidase could render a diseases-susceptible host more tolerant (Authié et al., 1992; Lalmanach et al., 2002). Using the idea of trypanotolerance attributed to antibody induction against *Tco*CATL, an anti-disease vaccine could be a possibility.

Besides *Tco*CATL and its affect on trypanotolerance, other parasitic cysteine peptidases are also virulence factors and have been shown to contribute to disease pathogenesis including migration and nutrition (Caffrey and Steverding, 2009). *Tcr*CATL, inhibited by the vinyl sulfone pseudopeptide inhibitor K11777, was effective in curing or alleviating *T. cruzi* infection in preclinical proof-of-concept studies and has now entered formal preclinical drug development (McKerrow et al., 2009). *Tcr*CATL shares 59% identity with *Tbb*CATL, 57% with *Tco*CATL and 45% with mammalian CATL (Eakin et al., 1992). Adding to the contribution to disease pathogenesis, *Tco*CATL degrades bovine fibrinogen and bovine serum albumin *in vitro* (Mbawa et al., 1992). *Tco*CATL is similar to *Tbb*CATL and *Tcr*CATL in that it is active at neutral pH allowing it to interact with host proteins in a physiological environment (Lalmanach et al., 2002). Furthermore, the recombinant catalytic domains of *Tvi*CATL (Vather, 2010) and *Tco*CATL (Pillay et al., 2010) are also active close to physiological pH. Although there are physiological natural cysteine peptidase inhibitors, incomplete enzyme inhibition with such inhibitors has been documented (Lonsdale-Eccles et al., 1995; Wiser et al., 1997). As seen on fibrinogen-containing SDS-PAGE gels, *Tbb*CATL and *Tco*CATL form active complexes of higher molecular weight when added to a rat kininogen-like cysteine peptidase inhibitor,

trypanopain modulator. Furthermore, the recombinant catalytic domain of *Tco*CATL forms a similar active complex when added to the cystatin inhibitor, stefin b, isolated from bovine liver (unpublished observation). Pyroglutamyl peptidase I (PGP), a cysteine peptidase with a slightly different catalytic triad (Glu instead of Asn); see below, is solely responsible for the reduced plasma half-life of thyrothrophin releasing hormone and partially responsible for the reduced half-life of gonadothrophin releasing hormone in *T. b. brucei* infected rodents (Morty et al., 2006).

Trypanosomal cysteine peptidases, like mammalian CATL, are synthesised as pre-pro-proteins (Cazzulo et al., 1989; Eakin et al., 1992; Sakanari et al., 1997). The catalytic domain, which contains the classical papain clan CA-like catalytic Cys, His and Asn triad, is linked to the 11-13 kDa C-terminal extension by a poly-threonine (*Tcr*CATL and *Tvi*CATL) or poly-proline (*Tbb*CATL and *Tco*CATL) tract. The C-terminal extension, for which a definitive biological function is not yet known (Caffrey et al., 2000), is unique to *Trypanosomatidae*. The overall 3-dimensional fold of cathepsin and papain-like cysteine peptidases is highly conserved. The prominent features of the structure consist of an L and R domain with a central α-helix and a β-barrel motif respectively (Turk et al., 2000).

Understanding the enzymatic as well as structural character of the cysteine peptidases that act as pathogenic factors in trypanosomosis will provide information that will facilitate better drug selection or improved development of a multi-component vaccine. Determining whether *T. vivax* CATL-like peptidases are pathogenic factors and similar in structure, function and character to other pathogenic trypanosomal CATL-like peptidases would be pivotal to the design of a general inhibitor for trypanosomal cysteine peptidases. CATL-like sequences from African and South American isolates of *T. vivax* have been evaluated for their suitability as genetic markers for population structure analysis and diagnosis (Cortez et al., 2009). The analysis of the CATL-like genes in that study did reveal that the *T. vivax* CATL-like sequences comprise the typical cysteine peptidase pre-, pro-, and central catalytic domain (with the Cys, His and Asn catalytic triad) as well as a C-terminal extension.

Enzyme characterisation of native *Tvi*CATL is the first step to deciphering how similar, in character, is *Tvi*CATL to other trypanosomal CATL-like cysteine peptidases.

Characterisation includes substrate preference, inhibitor preference, pH stability and the ability to degrade physiological proteins.

The functionally active full-length native *Tvi*CATL was purified from *T. vivax* strain Y486 parasites that had been grown in BALB/c mice. *Tvi*CATL as the case with Clan CA, family C1 peptidases showed a general sensitivity to E-64 and cystatin and has a substrate specificity defined by the S_2 pocket. Leu was preferred in the P_2 position and basic and non-bulky hydrophobic residues were accepted in the P_1 and P_3 positions respectively. Similarly, native *Tco*CATL (Mbawa et al., 1992), *Tbb*CATL (Troeberg et al., 1996) and *Tcr*CATL (Chagas et al., 1997) have a distinct preference for a basic amino acid (Arg of Lys) in the P_1 position. This is also a feature common to many mammalian and protozoan cysteine peptidases (Mbawa et al., 1992). Furthermore, the recombinant catalytic domain of *Tvi*CATL (Vather, 2010) and *Tco*CATL (Pillay et al., 2010) have similar amino acid preferences for the P_1 position. Native *Tvi*CATL favours a leucyl or phenylalanyl and not an arginyl residue in the P_2 position. Similar results were reported for the recombinant catalytic domains of *Tvi*CATL (Vather, 2010) and *Tco*CATL (Pillay et al., 2010), native *Tco*CATL (Mbawa et al., 1992), native *Tbb*CATL (Troeberg et al., 1996) and recombinant *Tbr*CATL (Caffrey et al., 2001).

*Tvi*CATL exhibited no activity towards the CATB-like substrate, Z-Arg-Arg-AMC but was able to hydrolyse Z-Phe-Arg-AMC, the CATL-like substrate. The hydrophobic pocket of the S_2 subsite of *Tcr*CATL is lined with residues Leu189, Met190, Ala258, Leu282 and Glu333. In the case of *Tco*CATL, a leucyl residue, which shifts on solvent exposure, is found at the bottom of the hydrophobic pocket instead of Glu333 (Lecaille et al., 2001). The substitution of a glutamyl by a leucyl residue does not allow for the accommodation of Arg in the P_2 position of *Tco*CATL. Sequence alignments also showed that *Tvi*CATL possesses a Leu instead of a Glu at this position, suggesting the inability of *Tvi*CATL to hydrolyse Z-Arg-Arg-AMC. In the case of *Tbr*CATL, an Ala occupies this position and Caffrey et al (2001) have demonstrated the incapability of *Tbr*CATL to degrade substrates with an Arg in the P_2 position.

*Tcr*CATL accepts Phe and Leu in the P_2 position with only a slight preference for Leu (Lima et al., 2001). However, recombinant *Tcr*CATL2 (a cysteine peptidase isoform) has a

marked preference for Leu over Phe (dos Reis et al., 2006). This suggests that the S_2 pocket of *Tcr*CATL2 might be more restricted than that of *Tcr*CATL. Structural studies of *Tcr*CATL have shown that, when in complex with irreversible inhibitors, Leu189, Ala258 and Leu282 line the S_2 pocket and readily accommodate the Phe residue at P_2, which interacts with the enzyme's backbone via the carbonyl of Gly188 (Brinen et al., 2000). A Ser substituted for Ala at position 258 and the proximal change of Met190 (found in *Tcr*CATL) to Ser190 in recombinant *Tcr*CATL2 may ultimately influence the enzyme's S_2 specificity (dos Reis et al., 2006). Based on a sequence alignment with *Tcr*CATL, residues lining the S_2 pocket of *Tvi*CATL are similar to those of *Tcr*CATL and do not have the relative amino acid substitutions found in *Tcr*CATL2. This supports the observation that *Tvi*CATL has only a slight preference for Leu over Phe in the S_2 position.

*Tvi*CATL was optimally active at pH 7.5. Slightly lower pH values have been reported for the optimal activity of native *Tcr*CATL (Chagas et al., 1997), *Tco*CATL (Chagas et al., 1997; Mbawa et al., 1992) and *Tbb*CATL (Troeberg et al., 1996). However these trypanosomal peptidases do possess activity at physiological pH. The only variation between native *Tvi*CATL and the recombinant catalytic domain of *Tvi*CATL (Vather, 2010) is the ability of the recombinant catalytic domain of *Tvi*CATL to remain active at acidic pH. Whether or not the C-terminal extension of *Tvi*CATL influences the activity at acidic pH values is still to be determined. It has, however, been shown that full length *Tbr*CATL as well as truncated *Tbr*CATL (without the C-terminal extension) have very similar pH profiles (Caffrey et al., 2001).

*Tvi*CATL, like other trypanosomal cysteine peptidases, was able to digest bovine fibrinogen. Additionally, *Tvi*CATL was effectively inhibited by chicken egg white cystatin, similar to native *Tco*CATL (Mbawa et al., 1992) and *Tcr*CATL (Lima et al., 2001). It was previously speculated that because of conserved sequences amongst cystatins, and the highly similar sequences of trypanosomal cysteine peptidases to those of the mammalian hosts, trypanosomal cysteine peptidases are likely to be inhibited by cystatins in the bloodstream of mammalian hosts (Troeberg et al., 1996). Peptidyl-diazomethane inhibitors such as, Z-RLVG-CHN_2, based on the cystatin-conserved QxVxG sequence, which forms part of the first β-hairpin loop and wedges into the papain-like cysteine peptidase substrate pocket (Auerswald et al., 1992), has been able to successfully

inhibit native *Tco*CATL and *Tcr*CATL (Chagas et al., 1997). However, the mammalian homologs such as CATB and CATL are inhibited to a greater extent with such peptidyl-diazomethane inhibitors. Attempts were made to overcome the limitations of cysteine peptidase inhibitors, specific for parasites, by labelling the N-terminus of substrates derived from the conserved residues from the N-terminal region of the cystatin superfamily. Labelling included biotin or using a spacer arm (6-aminohexanoic acid, Ahx) to increase the distance between the biotin and the peptidyl core of the inhibitor. Lalmanach et al. (1996) investigated the underlying premise that cysteine peptidases of the mammalian host and parasite might have different active site topologies that could favour or restrict the access of the biotinylated inhibitor to the active site cleft. This idea stemmed from the fact that the N-terminal, or Cbz group, of fluoromethylketone inhibitors protrudes from the active site of *Tcr*CATL into the solvent, making no contacts with the enzyme (McGrath et al., 1995). Biotinylated methylketone inhibitors have shown more specificity for *Tcr*CATL than mammalian CATL. The hydrophobic nature of the P_2-binding subsite was further confirmed by the ability of both chloromethyl ketones, Z-Gly-Leu-Phe-CMK and H-D-Val-Phe-Lys-CMK, to inhibit native *Tvi*CATL. As has been previously demonstrated for native *Tco*CATL, Z-Leu-Met-CHN_2 and Z-Leu-Lys-CHN_2 hydrophobic groups in P_2 (Leu) are very potent inhibitors (Mbawa et al., 1992). Similarly, Z-Phe-Phe-CMK and Z-Phe-Ala-FMK were potent inhibitors of *Tbb*CATL (Troeberg et al., 1999). Furthermore, K11777, the vinyl sulfone with a homophenyl group in P_1 and a phenyl group in the P_2 position was the most effective trypanocidal agent for cultured *T. b. brucei* and purified *Tbb*CATL (Troeberg et al., 1999). As mentioned earlier, K11777 has entered preclinical drug development investigations for curing *T. cruzi* infection.

It is important to know if the characteristics of mammalian cysteine peptidases differ from those of parasite homologs. Optimising the interaction of the active site of *Tvi*CATL with natural amino acid substrates has made considerable contributions to the development of inhibitors. Determining further structural and functional characteristics as well as whether *Tvi*CATL, like the *T. congolense*, *T.cruzi* and *T. b. brucei* homologs, *Tco*CATL, *Tcr*CATL and *Tbb*CATL respectively, acts as a pathogenic factor, would be essential for the designing of specific chemotherapeutic agents. However, phylogenetic analysis of sequences from East and West African *T. vivax* isolates, corresponding to CATL-like catalytic domains, revealed substantial polymorphism (Cortez et al., 2009). Furthermore, a

preliminary analysis of the organisation of CATL genes in the *T. vivax* genome showed that there is an array of 11 CATL-like genes organised in tandem array (Cortez et al., 2009). This tandem array for CATL genes is also found in the *T. congolense* (Kakundi, 2009), *T. cruzi* (Lima et al., 1994), *T. b. rhodesiense* (Mottram et al., 1989) and *Leishmania mexicana* genome (Mottram et al., 1997). As the characterisation of genes encoding such peptidases in *T. vivax* is the first step towards peptidase characterisation, which might contribute to the definition of targets for chemotherapy and/or vaccines, it is important that the organisation of the CATL genes is known. If *Tvi*CATL is proven to be pathogenic, then the other CATL-like peptidases encoded for in the *T. vivax* genome need be taken into consideration. These other CATL-like peptidases will need to be characterised and their structures determined. One of the most recurring substitutions of the *Tco*CATL isoforms is that of the active site Cys and His with Ser and Tyr respectively (Pillay et al., 2010). Although two *Tco*CATL isoforms have been shown to be active against classical cysteine peptidase substrates and inhibited by E-64, they differ slightly from one another and from the recombinant catalytic domain of *Tco*CATL in substrate preferences (Pillay et al., 2010). As previously mentioned, *Tcr*CATL2 does differ substantially from *Tcr*CATL regarding the S_2 subsite (Lima et al., 1994). Although methylketone inhibitors were able to inhibit purified native *Tvi*CATL, the ability of such inhibitors to effectively reduce *T. vivax* numbers in culture should also be tested. Furthermore, diazomethylketone inhibitors are relatively non-selective inhibitors and this to validate the target of methlyketone inhibitors and the biological role of *Tvi*CATL, RNAi on *Tvi*CATL should also be done to see whether there is a significantly different phenotypic feature in cultured *T. vivax*.

To date, *Tcr*CATL (Eakin et al., 1993; McGrath et al., 1995) and *Tbr*CATL (Kerr et al., 2009) are the only trypanosomal CATL-like cysteine peptidases that have been crystallised and their structures solved. Both structures follow a common two-domain fold of the papain superfamily. McGrath et al. (1995) reported that the active site of *Tcr*CATL extends over seven substrate-binding sites, four of which (S_1-S_4) are located on the acyl side of the cleaved bond and three (S_1'-S_3') on the amino side. The crystal structure of recombinant *Tcr*CATL in complex with K11017 (Mu-Leu-Hph-VSPh) shows that the hydrophobic residue in the P_2 position is able to make contact with the S_2 subsite of *Tcr*CATL as the S_2 subsite is lined with non-polar residues (Bryant et al., 2009; Kerr et al.,

2009). The advantage of structural analysis has allowed for the directed design of synthetic peptidase inhibitors.

The amino acid residues Lys39 and Glu44 have been shown to be involved in *Tco*CATL dimerisation. At acidic pH, the side chains of Lys39 and His43 are positively charged, whereas Glu44 is negatively charged and as a result, a salt bridge could be formed between His43 and Glu44, forming an α-helix (Ndlovu, 2010). However at physiological pH the His43 is deprotonated resulting in a salt bridge between Lys39 and Glu44, thus forming a β-turn, giving the protein dimeric conformation (Ndlovu, 2010). Mutating Lys39 and Glu44 to Phe39 and Pro44, prevents a salt bridge formation and therefore making *Tco*CATL exist as a monomer. Furthermore as previously speculated, mutating His43 to Trp43 forces the protein to exist as a dimer (Ndlovu, 2010). Consequently, structural determination is required to understand this dimerisation mechanism so that protective epitopes may be mapped. This will be valuable information needed for the development of an effective anti-disease vaccine against bovine trypanosomosis. Furthermore, details about the structure-function of *Tco*CATL will be of significance in the design of chemotherapeutic agents.

The last part of the study involved expression and purification of the recombinant catalytic domain of *Tvi*CATL and *Tco*CATL, *Tco*CATL dimerisation mutants and the full length PGP for crystallisation conditions screening. Using the Rigaku robotic crystallisation suite, purified proteins were screened with manual screens designed using variations of crystallising reagent conditions that have been used to previously crystallise related proteins. Furthermore, commercially available screens, with a large variation of crystallising reagents, were also used to screen the recombinant proteins. Crystals grown in the commercially available screens were used for X-ray diffraction.

The catalytic domain of *Tco*CATL and *Tvi*CATL were recombinantly expressed in *P. pastoris* and purified to homogeneity. PGP from *T. congolense* was recombinantly expressed in *E. coli* BL21 (DE3) cells and also purified to homogeneity. Purified cysteine peptidases along with previously purified *Tco*CATL dimerisation mutants, *Tco*CATL (H43W), *Tco*CATL (C25A), which contains the proenzyme, and *Tco*CATL (K39F; E44P) were screened for crystallisation conditions using the Rigaku robotic crystallisation suite.

Manual screens involved recording conditions from related proteins such as native papain (Kamphuis et al., 1984), native papain in complex with a recombinant cysteine peptidase inhibitor derived from *T. b. brucei* (Alphey and Hunter, 2006), native papain-E-64 complex (Varughese et al., 1989), human CATL-E-64 complex (Fujishima et al., 1997), recombinant *Tcr*CATL in complex with Z-Phe-Arg-CH_2F (McGrath et al., 1995) and K11777 (Kerr et al., 2009), recombinant *Tbr*CATL (without the C-terminal extension) in complex with K11777 and falcipain-3 (cysteine peptidase from *Plasmodium falciparum*) in complex with K11017 (Kerr et al., 2009). CrysPred predictor (Kantardjieff and Rupp, 2004) was also used to explore possible pH values for crystallisation, based on the proteins' pI values. Many of the results from the manual screen showed 'phase separation.' In a numerical scoring system, 'phase separation' scores five out of a possible ten (Zeelen, 2009).

Whilst manual optimisation screening is still underway, the commercial screens, JSCG+ and PACT, were used as another option. The two hits, using the PACT screen, found for *Tco*CATL (K39F; E44P) were further optimised. The crystals formed were one-dimensional needle-like and in order to get 'better' crystals with more dimension, conditions such as protein and crystallant concentration as well as buffer pH were varied. Analysis of a *Tco*CATL (K39F; E44P) crystal by X-ray diffraction confirmed that the crystal was protein however, the diffraction pattern was very weak and required further optimisation of crystal conditions to improve crystal quality. Whether or not the mutations of amino acid residues thought to be involved in the dimerisation, facilitated the crystallisation of *Tco*CATL (K39F; E44P) is still to be determined. However, *Tcr*CATL, which is always a monomer, has been successfully crystallised (Eakin et al., 1993; McGrath et al., 1995). Nonetheless, it has been demonstrated that the dimerisation mechanism of *Tco*CATL does involve more amino acid residues (Ndlovu, 2010). Most of the crystals of *Tco*CATL (K39F; E44P) were overnucleated. This was evident from the shower of needle-like one-dimensional crystals formed in crystallisation drops. Higher levels of supersaturation are required for crystallisation nuclei to form than for subsequent crystal growth (Luft and DeTitta, 1999). It has been shown that seeding, through either streak seeding (Stura and Wilson, 1991) or microseeding (Luft and DeTitta, 1999) is the most popular way for uncoupling nucleation, formed in the labile zone of the phase diagram (Fig. 1.9, chapter 1), from crystal growth, formed in metastable zone. Seeding techniques are still being pursued for crystallisation of the recombinant proteins.

Although all proteins were concentrated to >5mg/ml, which is sufficient for crystallisation, the inability of *Tvi*CATL and *Tco*CATL (C25A) to crystallise could have been attributed to glycosylation. Glycosylation has been said to interfere with crystallisation as the flexible and heterogeneous carbohydrate moieties can account for a significant fraction of the surface area of the protein, interfering with the crystallisation process (Derewenda, 2010). However, proteinase A of *Saccharomyces cerevisae* (Li et al., 2007) and HIV gp120 envelope protein (Kwong et al., 1998) have been crystallised in spite of the glycosylation. In the case of *Tbr*CATL crystallisation, however, the S172A mutation was made to remove the O-linked glycosylation site (Caffrey et al., 2001).

Although *Tco*CATL has pathogenic roles and provides insight into the better immune response of trypanotolerant cattle and hope for an anti-disease vaccine, trypanotolerance has been suggested to be multi-factorial, involving both non-immunological and immunological mechanisms (Authié, 1994). Furthermore, as previously mentioned, like that of CATL genes of *T. cruzi* (Lima et al., 1994) *T. b. rhodesiense* (Mottram et al., 1989), *T. vivax* (Cortez et al., 2009) and *L. mexicana* genome (Mottram et al., 1997), *Tco*CATL is organised in the genome of the parasite as a tandem array of multiple variants (Kakundi, 2009). Additionally, two of the variants of *Tco*CATL have been characterised and differ in substrate preference compared to *Tco*CATL (Pillay et al., 2010). Although the variants, like that of *Tco*CATL, prefer a hydrophobic amino acid, such as Leu or Phe, in the P_2 position, unlike *Tco*CATL, a small neutral amino acid, rather than a bulky cyclic amino acid such as proline, is required for the P_3 position (Pillay et al., 2010). As a result of these slight substrate preference differences, structural analysis is required as the substrate binding pockets may be different in size, character and geometry. These differences would be needed to be taken into consideration for proper drug and vaccine design. This unveils future work which would require structural studies on the variants of *Tco*CATL.

The initial studies on *Tvi*CATL were important preliminary studies to learn more about the relationship of the major cysteine peptidase of *T. vivax* with other trypanosomal cysteine peptidases, especially since to date, no such study has been done. Provided that *Tvi*CATL does play a role in pathogenesis, the commonality of enzymatic characteristics between *Tvi*CATL and those of *Tco*CATL, *Tbb*CATL, *Tbr*CATL and *Tcr*CATL does give hope for the design of a broader based drug. The crystal structure of *Tco*CATL, with some of the residues involved in dimerisation mutated, will be pivotal in understanding the

dimerisation model. This insight will be valuable towards therapies such as vaccine design and chemotherapeutics. Understanding the structure and function of the CATL-like cysteine peptidases of African trypanosomes is of major importance and significance to the control of African animal trypanosomosis.

REFERENCES

Abdulla, M. H., Lim, K. C., Sajid, M., McKerrow, J. H. and Caffrey, C. R. (2007). Schistosomiasis mansoni: novel chemotherapy using a cysteine protease inhibitor. *PLoS Med* **4**, e14.

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

Aerts, D., Truc, P., Penchenier, L., Claes, Y. and Le Ray, D. (1992). A kit for in vitro isolation of trypanosomes in the field: first trial with sleeping sickness patients in the Congo Republic. *Trans R Soc Trop Med Hyg* **86**, 394-395.

Ainanshe, O. A., Jennings, F. W. and Holmes, P. H. (1992). Isolation of drug-resistant strains of *Trypanosoma congolense* from the lower Shabelle region of southern Somalia. *Trop Anim Health Prod* **24**, 65-73.

Akol, M. N., Olaho-Mukani, W., Odiit, M., Enyaru, J. C., Matovu, E., Magona, J. and Okitoi, N. D. (1999). Trypanosomosis agglutination card test for *Trypanosoma brucei rhodesiense* sleeping sickness. *East Afr Med J* **76**, 38-41.

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

Aleshin, A. E., Hoffman, C., Firsov, L. M. and Honzatko, R. B. (1994). Refined crystal structures of glucoamylase from *Aspergillus awamori* var. X100. *J Mol Biol* **238**, 575-91.

Allsopp, R. and Hursey, B. H. (2004). Insecticidal control of tsetse. In *The Trypanosomiasis*: Oxfordshire: CABI Publishing.

Alphey, M. S. and Hunter, W. N. (2006). High-resolution complex of papain with remnants of a cysteine protease inhibitor derived from *Trypanosoma brucei*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**, 504-508.

Alvarez, V., Parussini, F., Aslund, L. and Cazzulo, J. J. (2002). Expression in insect cells of active mature cruzipain from *Trypanosoma cruzi*, containing its C-terminal domain. *Protein Expr Purif* **26**, 467-475.

Andrianarivo, A. G., Muiya, P., Opollo, M. and Logan-Henfrey, L. L. (1995). *Trypanosoma congolense*: comparative effects of a primary infection on bone marrow progenitor cells from N'Dama and Boran cattle. *Exp Parasitol* **80**, 407-418.

Aparicio, I. M., Scharfstein, J. and Lima, A. P. (2004). A new cruzipain-mediated pathway of human cell invasion by *Trypanosoma cruzi* requires trypomastigote membranes. *Infect Immun* **72**, 5892-5902.

Ashall, F., Angliker, H. and Shaw, E. (1990). Lysis of trypanosomes by peptidyl fluoromethyl ketones. *Biochem Biophys Res Commun* **170**, 923-929.

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

Auerswald, E. A., Genenger, G., Assfalg-Machleidt, I., Machleidt, W., Engh, R. A. and Fritz, H. (1992). Recombinant chicken egg white cystatin variants of the QLVSG region. *Eur J Biochem* **209**, 837-845.

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

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

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

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

Bajyana Songa, E. and Hamers, R. (1988). A card agglutination test (CATT) for veterinary use based on an early VAT RoTat 1/2 of *Trypanosoma evansi*. *Ann Soc Belg Med Trop* **68**, 233-240.

Balana-Fouce, R., Garzon Pulido, T., Ordonez-Escudero, D. and Garrido-Pertierra, A. (1986). Inhibition of diamine oxidase and S-adenosylmethionine decarboxylase by diminacene aceturate (berenil). *Biochem Pharmacol* **35**, 1597-1600.

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

Barrett, A. J. and Rawlings, N. D. (2004). Introduction: The clans and families of cysteine peptidases. In *Handbook of proteolytic enzymes*. London: Elsevier.

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.

Benvenuti, M. and Mangani, S. (2007). Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. *Nat Protoc* **2**, 1633-1651.

Bergfors, T. M. (1999). Dynamic light scattering. In *Protein Crystallisation: Stratergies, Techniques and Tips*. La Jolla: International University Line.

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. et al. (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science* **309**, 416-422.

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

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

Boibessot, I., Turner, C. M., Watson, D. G., Goldie, E., Connel, G., McIntosh, A., Grant, M. H. and Skellern, G. G. (2002). Metabolism and distribution of phenanthridine trypanocides in *Trypanosoma brucei*. *Acta Trop* **84**, 219-228.

Boid, R., Jones, T. W. and Payne, R. C. (1989). Malic enzyme type VII isoenzyme as an indicator of suramin resistance in *Trypanosoma evansi*. *Exp Parasitol* **69**, 317-323.

Bolanos-Garcia, V. M. and Chayen, N. E. (2009). New directions in conventional methods of protein crystallization. *Prog Biophys Mol Biol* **101**, 3-12.

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

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

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

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

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

Bray, P. G., Barrett, M. P., Ward, S. A. and de Koning, H. P. (2003). Pentamidine uptake and resistance in pathogenic protozoa: past, present and future. *Trends Parasitol* **19**, 232-239.

Brinen, L. S., Hansell, E., Cheng, J., Roush, W. R., McKerrow, J. H. and Fletterick, R. J. (2000). A target within the target: probing cruzain's P1' site to define structural determinants for the Chagas' disease protease. *Structure* **8**, 831-840.

Brömme, D., Li, Z., Barnes, M. and Mehler, E. (1999). Human cathepsin V functional expression, tissue distribution, electrostatic surface potential, enzymatic characterization, and chromosomal localization. *Biochemistry* **38**, 2377-2385.

Bruning, A. (1996). Equine piroplasmosis an update on diagnosis, treatment and prevention. *Br Vet J* **152**, 139-151.

Bryant, C., Kerr, I. D., Debnath, M., Ang, K. K., Ratnam, J., Ferreira, R. S., Jaishankar, P., Zhao, D., Arkin, M. R., McKerrow, J. H., Brinen, L. S. and Renslo, A. R. (2009). Novel non-peptidic vinylsulfones targeting the S2 and S3 subsites of parasite cysteine proteases. *Bioorg Med Chem Lett* **19**, 6218-6221.

Buza, J. and Naessens, J. (1999). Trypanosome non-specific IgM antibodies detected in serum of *Trypanosoma congolense*-infected cattle are polyreactive. *Vet Immunol Immunopathol* **69**, 1-9.

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

Caffrey, C. R., Scory, S. and Steverding, D. (2000). Cysteine proteinases of trypanosome parasites: novel targets for chemotherapy. *Curr Drug Targets* **1**, 155-162.

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

Caffrey, C. R. and Steverding, D. (2009). Kinetoplastid papain-like cysteine peptidases. *Mol Biochem Parasitol* **167**, 12-9.

Campbell, G. A. and West, T. F. (1952). DDT and newer persistent insecticides. New York: Chemcial Pub. Co.

Cazzulo, J. J., Couso, R., Raimondi, A., Wernstedt, C. and Hellman, U. (1989). Further characterization and partial amino acid sequence of a cysteine proteinase from *Trypanosoma cruzi*. *Mol Biochem Parasitol* **33**, 33-41.

Cazzulo, J. J., Cazzulo Franke, M. C., Martinez, J. and Franke de Cazzulo, B. M. (1990). Some kinetic properties of a cysteine proteinase (cruzipain) from Trypanosoma cruzi. *Biochim Biophys Acta* **1037**, 186-191.

Cazzulo, J. J., Stoka, V. and Turk, V. (1997). Cruzipain, the major cysteine proteinase from the protozoan parasite *Trypanosoma cruzi*. *Biol Chem* **378**, 1-10.

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

Chan, V. J., Selzer, P. M., McKerrow, J. H. and Sakanari, J. A. (1999). Expression and alteration of the S2 subsite of the *Leishmania major* cathepsin B-like cysteine protease. *Biochem J* **340 (Pt 1)**, 113-117.

Chang, J., Jallouli, Y., Kroubi, M., Yuan, X. B., Feng, W., Kang, C. S., Pu, P. Y. and Betbeder, D. (2009). Characterization of endocytosis of transferrin-coated PLGA nanoparticles by the blood-brain barrier. *Int J Pharm* **379**, 285-292.

Chen, Y. T., Lira, R., Hansell, E., McKerrow, J. H. and Roush, W. R. (2008). Synthesis of macrocyclic trypanosomal cysteine protease inhibitors. *Bioorg Med Chem Lett* **18**, 5860-5863.

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

Choe, Y., Leonetti, F., Greenbaum, D. C., Lecaille, F., Bogyo, M., Bromme, D., Ellman, J. A. and Craik, C. S. (2006). Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. *J Biol Chem* **281**, 12824-12832.

Compton, S. J. and Jones, C. G. (1985). Mechanism of dye response and interference in the Bradford protein assay. *Anal Biochem* **151**, 369-374.

Cortez, A. P., Rodrigues, A. C., Garcia, H. A., Neves, L., Batista, J. S., Bengaly, Z., Paiva, F. and Teixeira, M. M. (2009). Cathepsin L-like genes of *Trypanosoma vivax* from Africa and South America-characterization, relationships and diagnostic implications. *Mol Cell Probes* **23**, 44-51.

Coulombe, R., Li, Y., Takebe, S., Menard, R., Mason, P., Mort, J. S. and Cygler, M. (1996). Crystallization and preliminary X-ray diffraction studies of human procathepsin L. *Proteins* **25**, 398-400.

Dame, J. B., Arnot, D. E., Bourke, P. F., Chakrabarti, D., Christodoulou, Z., Coppel, R. L., Cowman, A. F., Craig, A. G., Fischer, K., Foster, J., Goodman, N., Hinterberg, K., Holder, A. A., Holt, D. C., Kemp, D. J., Lanzer, M., Lim, A., Newbold, C. I., Ravetch, J. V., Reddy, G. R., Rubio, J., Schuster, S. M., Su, X. Z., Thompson, J. K., Werner, E. B. and et al. (1996). Current status of the Plasmodium falciparum genome project. *Mol Biochem Parasitol* **79**, 1-12.

de Koning, H. P., Anderson, L. F., Stewart, M., Burchmore, R. J., Wallace, L. J. and Barrett, M. P. (2004). The trypanocide diminazene aceturate is accumulated predominantly through the TbAT1 purine transporter: additional insights on diamidine resistance in african trypanosomes. *Antimicrob Agents Chemother* **48**, 1515-1519.

Del Nery, E., Juliano, M. A., Lima, A. P., Scharfstein, J. and Juliano, L. (1997). Kininogenase activity by the major cysteinyl proteinase (cruzipain) from *Trypanosoma cruzi*. *J Biol Chem* **272**, 25713-25718.

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

Delespaux, V., Dinka, H., Masumu, J., Van den Bossche, P. and Geerts, S. (2008). Five-fold increase in Trypanosoma congolense isolates resistant to diminazene aceturate over a seven-year period in Eastern Zambia. *Drug Resist Updat* **11**, 205-209.

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

Derewenda, Z. S. (2004). The use of recombinant methods and molecular engineering in protein crystallization. *Methods* **34**, 354-363.

Derewenda, Z. S. (2010). Application of protein engineering to enhance crystallizability and improve crystal properties. *Acta Crystallogr D Biol Crystallogr* **66**, 604-615.

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

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

dos Reis, F. C., Judice, W. A., Juliano, M. A., Juliano, L., Scharfstein, J. and Lima, A. P. (2006). The substrate specificity of cruzipain 2, a cysteine protease isoform from Trypanosoma cruzi. *FEMS Microbiol Lett* **259**, 215-220.

Dua, M., Raphael, P., Sijwali, P. S., Rosenthal, P. J. and Hanspal, M. (2001). Recombinant falcipain-2 cleaves erythrocyte membrane ankyrin and protein 4.1. *Mol Biochem Parasitol* **116**, 95-99.

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

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

Eisler, M. C., Lessard, P., Masake, R. A., Moloo, S. K. and Peregrine, A. S. (1998). Sensitivity and specificity of antigen-capture ELISAs for diagnosis of *Trypanosoma congolense* and *Trypanosoma vivax* infections in cattle. *Vet Parasitol* **79**, 187-201.

Eisler, M. C., Torr, S. J., Coleman, P. G., Machila, N. and Morton, J. F. (2003). Integrated control of vector-borne diseases of livestock--pyrethroids: panacea or poison? *Trends Parasitol* **19**, 341-345.

Eisler, M. C., Dwinger, R. H., Majiwa, P. A. O. and Picozzi, K. (2004). Diagnosis and epidemiology of African animal trypanosomosis. In *The Trypanosomiasis*: Oxfordshire: CABI Publishing.

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

Elliott, M., Farnham, A. W., Janes, N. F., Needham, P. H. and Pulman, D. A. (1974). Synthetic insecticide with a new order of activity. *Nature* **248**, 710-771.

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

Engel, J. C., Doyle, P. S., Hsieh, I. and McKerrow, J. H. (1998a). Cysteine protease inhibitors cure an experimental *Trypanosoma cruzi* infection. *J Exp Med* **188**, 725-734.

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

Ewing, F. L., Forsythe, E. L., Van de Woerd, M. and Pusey, M. L. (1996). Effects of purification of the crystallization of lysozyme. *J Cryst Growth* **160**, 389-397.

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.

Franke de Cazzulo, B. M., Martinez, J., North, M. J., Coombs, G. H. and Cazzulo, J. J. (1994). Effects of proteinase inhibitors on the growth and differentiation of *Trypanosoma cruzi*. *FEMS Microbiol Lett* **124**, 81-86.

Fujinaga M., Cherney M.M, Oyama H., Oda K., James M.N.G. (2004) The molecular structure and catalytic mechanism of a novel carboxyl peptidase from *Scytalidium lignicolum. Proc. Natl Acad.* **101**,3364-3369.

Fujishima, A., Imai, Y., Nomura, T., Fujisawa, Y., Yamamoto, Y. and Sugawara, T. (1997). The crystal structure of human cathepsin L complexed with E-64. *FEBS Lett* **407**, 47-50.

Gibson, W. and Stevens, J. (1999). Genetic exchange in the *trypanosomatidae*. *Adv Parasitol* **43**, 1-46.

Gilliland, G. L. and Ladner, J. E. (1996). Crystallization of biological macromolecules for X-ray diffraction studies. *Curr Opin Struct Biol* **6**, 595-603.

Gillmor, S. A., Craik, C. S. and Fletterick, R. J. (1997). Structural determinants of specificity in the cysteine protease cruzain. *Protein Sci* **6**, 1603-1611.

Guncar, G., Pungercic, G., Klemencic, I., Turk, V. and Turk, D. (1999). Crystal structure of MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S. *EMBO J* **18**, 793-803.

Guo, Y. L., Kurz, U., Schultz, J. E., Lim, C. C., Wiederanders, B. and Schilling, K. (2000). The alpha1/2 helical backbone of the prodomains defines the intrinsic inhibitory specificity in the cathepsin L-like cysteine protease subfamily. *FEBS Lett* **469**, 203-207.

Hamilton, P. B., Stevens, J. R., Gaunt, M. W., Gidley, J. and Gibson, W. C. (2004). Trypanosomes are monophyletic: evidence from genes for glyceraldehyde phosphate dehydrogenase and small subunit ribosomal RNA. *Int J Parasitol* **34**, 1393-1404.

Harris, J. L., Backes, B. J., Leonetti, F., Mahrus, S., Ellman, J. A. and Craik, C. S. (2000). Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. *Proc Natl Acad Sci U S A* **97**, 7754-7759.

Helms, M. J., Ambit, A., Appleton, P., Tetley, L., Coombs, G. H. and Mottram, J. C. (2006). Bloodstream form *Trypanosoma brucei* depend upon multiple metacaspases associated with RAB11-positive endosomes. *J Cell Sci* **119**, 1105-1117.

Hertz-Fowler, C., Figueiredo, L. M., Quail, M. A., Becker, M., Jackson, A., Bason, N., Brooks, K., Churcher, C., Fahkro, S., Goodhead, I., Heath, P., Kartvelishvili, M., Mungall, K., Harris, D., Hauser, H., Sanders, M., Saunders, D., Seeger, K., Sharp, S., Taylor, J. E., Walker, D., White, B., Young, R., Cross, G. A., Rudenko, G., Barry, J. D., Louis, E. J. and Berriman, M. (2008). Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS One* **3**, e3527.

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

Hogg, T., Nagarajan, K., Herzberg, S., Chen, L., Shen, X., Jiang, H., Wecke, M., Blohmke, C., Hilgenfeld, R. and Schmidt, C. L. (2006). Structural and functional characterization of Falcipain-2, a hemoglobinase from the malarial parasite *Plasmodium falciparum*. *J Biol Chem* **281**, 25425-25437.

Holmes, P. H., Eisler, M. C. and Geerts, S. (2004). Current chemotherapy of animal trypanososis. In *The Trypanosomosis*: Oxfordshire: CABI Publishing.

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

Huang, L., Brinen, L. S. and Ellman, J. A. (2003). Crystal structures of reversible ketone-based inhibitors of the cysteine protease cruzain. *Bioorg Med Chem* **11**, 21-9.

Judice, W. A., Cezari, M. H., Lima, A. P., Scharfstein, J., Chagas, J. R., Tersariol, I. L., Juliano, M. A. and Juliano, L. (2001). Comparison of the specificity, stability and individual rate constants with respective activation parameters for the peptidase activity of cruzipain and its recombinant form, cruzain, from Trypanosoma cruzi. *Eur J Biochem* **268**, 6578-6586.

Kabayo, J. P. (2002). Aiming to eliminate tsetse from Africa. *Trends Parasitol* **18**, 473-5.

Kakundi, E. M. (2009). Molecular analysis of the congopain gene family. MSc dissertation. University of KwaZulu-Natal (Pietermaritzburg campus).

Kamphuis, I. G., Kalk, K. H., Swarte, M. B. and Drenth, J. (1984). Structure of papain refined at 1.65 A resolution. *J Mol Biol* **179**, 233-256.

Kamphuis, I. G., Drenth, J. and Baker, E. N. (1985). Thiol proteases. Comparative studies based on the high-resolution structures of papain and actinidin, and on amino acid sequence information for cathepsins B and H, and stem bromelain. *J Mol Biol* **182**, 317-329.

Kantardjieff, K. A. and Rupp, B. (2004). Protein isoelectric point as a predictor for increased crystallization screening efficiency. *Bioinformatics* **20**, 2162-2168.

Karrer, K. M., Peiffer, S. L. and DiTomas, M. E. (1993). Two distinct gene subfamilies within the family of cysteine protease genes. *Proc Natl Acad Sci U S A* **90**, 3063-3067.

Karvonen, E., Kauppinen, L., Partanen, T. and Poso, H. (1985). Irreversible inhibition of putrescine-stimulated S-adenosyl-L-methionine decarboxylase by berenil and pentamidine. *Biochem J* **231**, 165-169.

Kerr, I. D., Lee, J. H., Farady, C. J., Marion, R., Rickert, M., Sajid, M., Pandey, K. C., Caffrey, C. R., Legac, J., Hansell, E., McKerrow, J. H., Craik, C. S., Rosenthal, P. J. and Brinen, L. S. (2009). Vinyl sulfones as antiparasitic agents and a structural basis for drug design. *J Biol Chem* **284**, 25697-25703.

Kristjanson, P. M., Swallow, B. M., Kruska, R. L. and Leeuw, P. N. D. (1999). Measuring the cost of African animal trypanosomosis, the potential benefits of control and returns to research. *Agric Sys*, 79-98.

Kroubi, M., Karembe, H. and Betbeder, D. (2011). Drug delivery systems in the treatment of African trypanosomiasis infections. *Expert Opin Drug Deliv* **8**, 735-747.

Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J. and Hendrickson, W. A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648-659.

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

Lalmanach, G., Mayer, R., Serveau, C., Scharfstein, J. and Gauthier, F. (1996). Biotin-labelled peptidyl diazomethane inhibitors derived from the substrate-like sequence of cystatin: targeting of the active site of cruzipain, the major cysteine proteinase of Trypanosoma cruzi. *Biochem J* **318 (Pt 2)**, 395-399.

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

Lanham, S. M. and Godfrey, D. G. (1970). Isolation of salivarian trypanosomes from man and other animals using DEAE-cellulose. *Experimental Parasitology* **28**, 521.

Laveissiere, C., Camara, M., Rayaisse, J. B., Salou, E., Kagbadouno, M. and Solano, P. (2011). Trapping tsetse flies on water. *Parasite* **18**, 141-144.

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

Lecaille, F., 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. *Chem Rev* **102**, 4459-4488.

Lejon, V., Rebeski, D. E., Ndao, M., Baelmans, R., Winger, E. M., Faye, D., Geerts, S. and Buscher, P. (2003). Performance of enzyme-linked immunosorbent assays for detection of antibodies against *T. congolense* and *T. vivax* in goats. *Vet Parasitol* **116**, 87-95.

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

Lima, A. P., Scharfstein, J., Storer, A. C. and Menard, R. (1992). Temperature-dependent substrate inhibition of the cysteine proteinase (GP57/51) from Trypanosoma cruzi. *Mol Biochem Parasitol* **56**, 335-338.

Lima, A. P., Tessier, D. C., Thomas, D. Y., Scharfstein, J., Storer, A. C. and Vernet, T. (1994). Identification of new cysteine protease gene isoforms in *Trypanosoma cruzi*. *Mol Biochem Parasitol* **67**, 333-338.

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. *Mol Biochem Parasitol* **114**, 41-52.

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

Lonsdale-Eccles, J. D. and Grab, D. J. (2002). Trypanosome hydrolases and the blood-brain barrier. *Trends Parasitol* **18**, 17-9.

Luckins, A. G. (1992). Trypanosomosis in small ruminants-a major constraint to livestock production? *Br Vet J* **148**, 471-473.

Luft, J. R., Albright, D. T., Baird, J. K. and DeTitta, G. T. (1996). The rate of water equilibration in vapor-diffusion crystallizations: dependence on the distance from the droplet to the reservoir. *Acta Crystallogr D Biol Crystallogr* **52**, 1098-1106.

Luft, J. R. and DeTitta, G. T. (1999). A method to produce microseed stock for use in the crystallization of biological macromolecules. *Acta Crystallogr D Biol Crystallogr* **55**, 988-993.

Luft, J. R. and DeTitta, G. T. (2009). Rational selection of crystallization techniques. In *Protein Crystallization, 2nd Edition.*, (ed. T. M. Bergfors). La Jolla: International University Line.

Lukes, J., Jirku, M., Dolezel, D., Kral'ova, I., Hollar, L. and Maslov, D. A. (1997). Analysis of ribosomal RNA genes suggests that trypanosomes are monophyletic. *J Mol Evol* **44**, 521-527.

Lun, Z. R., Min, Z. P., Huang, D., Liang, J. X., Yang, X. F. and Huang, Y. T. (1991). Cymelarsan in the treatment of buffaloes naturally infected with *Trypanosoma evansi* in south China. *Acta Trop* **49**, 233-236.

Magez, S., Caljon, G., Tran, T., Stijlemans, B. and Radwanska, M. (2009). Current status of vaccination against African trypanosomiasis. *Parasitology* **137**, 2017-2027.

Maillard, J. C., Berthier, D., Thevenon, S., Quere, R., Piquemal, D., Manchon, L. and Marti, J. (2004). Use of the Serial Analysis of Gene Expression (SAGE) method in veterinary research: A concrete application in the study of the bovine trypanotolerance genetic control. *Ann N Y Acad Sci* **1026**, 171-182.

Maillard, J. C., Berthier, D., Thevenon, S., Piquemal, D., Chantal, I. and Marti, J. (2005). Efficiency and limits of the Serial Analysis of Gene Expression (SAGE) method: discussions based on first results in bovine trypanotolerance. *Vet Immunol Immunopathol* **108**, 59-69.

Marcello, L. and Barry, J. D. (2007). Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Res* **17**, 1344-1352.

Mark, B. L., Mahuran, D. J., Cherney, M. M., Zhao, D., Knapp, S. and James, M. N. (2003). Crystal structure of human beta-hexosaminidase B: understanding the molecular basis of Sandhoff and Tay-Sachs disease. *J Mol Biol* **327**, 1093-1109.

Masake, R. A. and Nantulya, V. M. (1991). Sensitivity of an antigen detection enzyme immunoassay for diagnosis of *Trypanosoma congolense* infections in goats and cattle. *J Parasitol* **77**, 231-236.

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

Maudlin, I., Turner, M. J., Dukes, P. and Miller, N. (1984). Maintenance of *Glossina morsitans morsitans* on antiserum to procyclic trypanosomes reduces infection rates with homologous and heterologous *Trypanosoma congolense* stocks. *Acta Trop* **41**, 253-257.

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

McCulloch, R. (2004). Antigenic variation in African trypanosomes: monitoring progress. *Trends Parasitol* **20**, 117-121.

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

McKerrow, J. H. (1999). Development of cysteine protease inhibitors as chemotherapy for parasitic diseases: insights on safety, target validation, and mechanism of action. *Int J Parasitol* **29**, 833-837.

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

McKerrow, J. H., Rosenthal, P. J., Swenerton, R. and Doyle, P. (2008). Development of protease inhibitors for protozoan infections. *Curr Opin Infect Dis* **21**, 668-72.

McKerrow, J. H., Doyle, P. S., Engel, J. C., Podust, L. M., Robertson, S. A., Ferreira, R., Saxton, T., Arkin, M., Kerr, I. D., Brinen, L. S. and Craik, C. S. (2009). Two approaches to discovering and developing new drugs for Chagas disease. *Mem Inst Oswaldo Cruz* **104 Suppl 1**, 263-269.

McNamara, J. J., Bailey, J. W., Smith, D. H., Wakhooli, S. and Godfrey, D. G. (1995). Isolation of *Trypanosoma brucei gambiense* from northern Uganda: evaluation of the kit for in vitro isolation (KIVI) in an epidemic focus. *Trans R Soc Trop Med Hyg* **89**, 388-389.

McPherson, A., Malkin, A. J. and Kuznetsov, Y. G. (1995). The science of macromolecular crystallization. *Structure* **3**, 759-768.

McPherson, A. (2009). Some words of advice from an old hand. In *Protein Crystallisation, 2nd edition*. La Jolla: International Univeristy Line.

Meade, K. G., O'Gorman, G. M., Hill, E. W., Narciandi, F., Agaba, M., Kemp, S. J., O'Farrelly, C. and MacHugh, D. E. (2009). Divergent antimicrobial peptide (AMP) and acute phase protein (APP) responses to *Trypanosoma congolense* infection in trypanotolerant and trypanosusceptible cattle. *Mol Immunol* **47**, 196-204.

Mendoza-Lopez, M. R., Becerril-Garcia, C., Fattel-Facenda, L. V., Avila-Gonzalez, L., Ruiz-Tachiquin, M. E., Ortega-Lopez, J. and Arroyo, R. (2000). CP30, a cysteine proteinase involved in *Trichomonas vaginalis* cytoadherence. *Infect Immun* **68**, 4907-4912.

Mkhize, P. P. (2003). Epitope mapping of a trypanosomal cysteine protease. MSc dissertation. University of KwaZulu-Natal (Pietermaritzburg campus).

Moreira, D., Lopez-Garcia, P. and Vickerman, K. (2004). An updated view of kinetoplastid phylogeny using environmental sequences and a closer outgroup: proposal for a new classification of the class Kinetoplastea. *Int J Syst Evol Microbiol* **54**, 1861-1875.

Mori, Y., Nagamine, K., Tomita, N. and Notomi, T. (2001). Detection of loop-mediated isothermal amplification reaction by turbidity derived from magnesium pyrophosphate formation. *Biochem Biophys Res Commun* **289**, 150-154.

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

Morrison, W. I., Black, S. J., Paris, J., Hinson, C. A. and Wells, P. W. (1982). Protective immunity and specificity of antibody responses elicited in cattle by irradiated *Trypanosoma brucei*. *Parasite Immunol* **4**, 395-407.

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

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

Mucache, H. (2011). *Trypanosoma congolense* pyroglutamyl peptidase: enzymatic characterisation and gene deletion. MSc dissertation. University of KwaZulu-Natal (Pietermaritzburg campus).

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

Murray, M., Morrison, W. I. and Whitelaw, D. D. (1982). Host susceptibility to African trypanosomiasis: trypanotolerance. *Adv Parasitol* **21**, 1-68.

Murray, M. and Dexter, T. M. (1988). Anaemia in bovine African trypanosomiasis. A review. *Acta Trop* **45**, 389-432.

Na, B. K., Bae, Y. A., Zo, Y. G., Choe, Y., Kim, S. H., Desai, P. V., Avery, M. A., Craik, C. S., Kim, T. S., Rosenthal, P. J. and Kong, Y. (2010). Biochemical properties of a novel cysteine protease of Plasmodium vivax, vivapain-4. *PLoS Negl Trop Dis* **4**, e849.

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*. *Vet Parasitol* **111**, 125-142.

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

Nantulya, V. M., Musoke, A. J., Rurangirwa, F. R., Saigar, N. and Minja, S. H. (1987). Monoclonal antibodies that distinguish *Trypanosoma congolense*, *T. vivax* and *T. brucei*. *Parasite Immunol* **9**, 421-431.

Nantulya, V. M., Lindqvist, K. J., Stevenson, P. and Mwangi, E. K. (1992). Application of a monoclonal antibody-based antigen detection enzyme-linked immunosorbent assay (antigen ELISA) for field diagnosis of bovine trypanosomiasis at Nguruman, Kenya. *Ann Trop Med Parasitol* **86**, 225-230.

Ndlovu, H. H. (2010). Structural studies aimed at improving the antigenicity of congopain. In *Biochemistry*,. Msc dissertation (ed. Pietermaritzburg: University of KwaZulu-Natal.

Ndoutamia, G., Moloo, S. K., Murphy, N. B. and Peregrine, A. S. (1993). Derivation and characterization of a quinapyramine-resistant clone of *Trypanosoma congolense*. *Antimicrob Agents Chemother* **37**, 1163-1166.

Nery, E. D., Juliano, M. A., Meldal, M., Svendsen, I., Scharfstein, J., Walmsley, A. and Juliano, L. (1997). Characterization of the substrate specificity of the major cysteine protease (cruzipain) from Trypanosoma cruzi using a portion-mixing combinatorial library and fluorogenic peptides. *Biochem J* **323 (Pt 2)**, 427-433.

Newman, J. (2009). Two approaches for intial screening: Evolution and Intelligent design. In Protein Crystallization, 2nd ed. La Jolla: International University Line.

Nissler, K., Kreusch, S., Rommerskirch, W., Strubel, W., Weber, E. and Wiederanders, B. (1998). Sorting of non-glycosylated human procathepsin S in mammalian cells. *Biol Chem* **379**, 219-224.

Noël, W., Hassanzadeh, G., Raes, G., Namangala, B., Daems, I., Brys, L., Brombacher, F., Baetselier, P. D. and Beschin, A. (2002). Infection stage-dependent modulation of macrophage activation in *Trypanosoma congolense*-resistant and -susceptible mice. *Infect Immun* **70**, 6180-6187.

Olaya, P. and Wasserman, M. (1991). Effect of calpain inhibitors on the invasion of human erythrocytes by the parasite *Plasmodium falciparum*. *Biochim Biophys Acta* **1096**, 217-221.

Ostermeier, C., Harrenga, A., Ermler, U. and Michel, H. (1997). Structure at 2.7 A resolution of the *Paracoccus denitrificans* two-subunit cytochrome c oxidase complexed with an antibody FV fragment. *Proc Natl Acad Sci U S A* **94**, 10547-10553.

Page, R., Grzechnik, S. K., Canaves, J. M., Spraggon, G., Kreusch, A., Kuhn, P., Stevens, R. C. and Lesley, S. A. (2003). Shotgun crystallization strategy for structural genomics: an optimized two-tiered crystallization screen against the *Thermotoga maritima* proteome. *Acta Crystallogr D Biol Crystallogr* **59**, 1028-1037.

Page, R. and Stevens, R. C. (2004). Crystallization data mining in structural genomics: using positive and negative results to optimize protein crystallization screens. *Methods* **34**, 373-389.

Palmer, J. T., Rasnick, D., Klaus, J. L. and Bromme, D. (1995). Vinyl sulfones as mechanism-based cysteine protease inhibitors. *J Med Chem* **38**, 3193-3196.

Pamer, E. G., Davis, C. E., Eakin, A. and So, M. (1990). Cloning and sequencing of the cysteine protease cDNA from Trypanosoma brucei rhodesiense. *Nucleic Acids Res* **18**, 6141.

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

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

Pinder, M., Libeau, G., Hirsch, W., Tamboura, I., Hauck-Bauer, R. and Roelants, G. E. (1984). Anti-trypanosome specific immune responses in bovids of differing susceptibility to African trypanosomiasis. *Immunology* **51**, 247-258.

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

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

Rae, P. F. and Luckins, A. G. (1984). Detection of circulating trypanosomal antigens by enzyme immunoassay. *Ann Trop Med Parasitol* **78**, 587-596.

Rautenberg, P., Schadler, R., Reinwald, E. and Risse, H. J. (1982). Study on a proteolytic enzyme from Trypansoma congolense. *Molecular Cellular Biochemstry* **47**, 151-159.

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

Rawlings, N. D. and Barrett, A. J. (1999). MEROPS: the peptidase database. *Nucleic Acids Res* **27**, 325-331.

Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. and Barrett, A. J. (2008). MEROPS: the peptidase database. *Nucleic Acids Res* **36**, D320-5.

Rawlings, N. D., Barrett, A. J. and Bateman, A. (2011). A seventh catalytic type of proteolytic enzyme**s.** *Biol Chem* **286,** 38321-38328.

Rebeski, D. E., Winger, E. M., Ouma, J. O., Kong Pages, S., Buscher, P., Sanogo, Y., Dwinger, R. H. and Crowther, J. R. (2001). Charting methods to monitor the operational performance of ELISA method for the detection of antibodies against trypanosomes. *Vet Parasitol* **96**, 11-50.

Reid, S. A. and Copeman, D. B. (2003). The development and validation of an antibody-ELISA to detect *Trypanosoma evansi* infection in cattle in Australia and Papua New Guinea. *Prev Vet Med* **61**, 195-208.

Reinert, K. E. (1999). DNA multimode interaction with berenil and pentamidine; double helix stiffening, unbending and bending. *J Biomol Struct Dyn* **17**, 311-331.

Robertson, C. D. and Coombs, G. H. (1990). Characterisation of three groups of cysteine proteinases in the amastigotes of *Leishmania mexicana mexicana*. *Mol Biochem Parasitol* **42**, 269-276.

Roditi, I. and Lehane, M. J. (2008). Interactions between trypanosomes and tsetse flies. *Curr Opin Microbiol* **11**, 345-351.

Rupp, B. and Wang, J. (2004). Predictive models for protein crystallization. *Methods* **34**, 390-407.

Rupp, B. (2009). Strategies for protein crystallisation screening. In *Protein Crystallization, 2nd edition,* (ed. T. M. Bergfor). La Jolla: International University Line.

Sajid, M. and McKerrow, J. H. (2002). Cysteine proteases of parasitic organisms. *Mol Biochem Parasitol* **120**, 1-21.

Sakanari, J. A., Nadler, S. A., Chan, V. J., Engel, J. C., Leptak, C. and Bouvier, J. (1997). *Leishmania major*: comparison of the cathepsin L- and B-like cysteine protease genes with those of other trypanosomatids. *Exp Parasitol* **85**, 63-76.

Salvesen, G. and Nagase, H. (1989). Inhibition of proteolytic enzymes. Oxford: Oxford University Press.

Schechter, I. and Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27**, 157-162.

Schilling, K., Pietschmann, S., Fehn, M., Wenz, I. and Wiederanders, B. (2001). Folding incompetence of cathepsin L-like cysteine proteases may be compensated by the highly conserved, domain-building N-terminal extension of the proregion. *Biol Chem* **382**, 859-865.

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

Shapiro, T. A. and Englund, P. T. (1990). Selective cleavage of kinetoplast DNA minicircles promoted by antitrypanosomal drugs. *Proc Natl Acad Sci U S A* **87**, 950-954.

Shenai, B. R., Sijwali, P. S., Singh, A. and Rosenthal, P. J. (2000). Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of *Plasmodium falciparum*. *J Biol Chem* **275**, 29000-29010.

Sileghem, M. and Naessens, J. (1995). Are CD8 T cells involved in control of African trypanosomiasis in a natural host environment? *Eur J Immunol* **25**, 1965-1971.

Simpson, A. G., Lukes, J. and Roger, A. J. (2002). The evolutionary history of kinetoplastids and their kinetoplasts. *Mol Biol Evol* **19**, 2071-2083.

Sinyangwe, L., Delespaux, V., Brandt, J., Geerts, S., Mubanga, J., Machila, N., Holmes, P. H. and Eisler, M. C. (2004). Trypanocidal drug resistance in eastern province of Zambia. *Vet Parasitol* **119**, 125-135.

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. and Klenk, D. C. (1985). Measurement of protein using bicinchoninic acid. *Anal Biochem* **150**, 76-85.

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

Stack, C. M., Caffrey, C. R., Donnelly, S. M., Seshaadri, A., Lowther, J., Tort, J. F., Collins, P. R., Robinson, M. W., Xu, W., McKerrow, J. H., Craik, C. S., Geiger, S. R., Marion, R., Brinen, L. S. and Dalton, J. P. (2008). Structural and functional relationships in the virulence-associated cathepsin L proteases of the parasitic liver fluke, *Fasciola hepatica*. *J Biol Chem* **283**, 9896-9908.

Stevens, J. R. and Brisse, S. (2004). Systematics of trypanosomes of medical and veterinary importance. In *The Trypanosomiases*: Oxfordshire: CABI Publishing.

Stevens, J. R. (2008). Kinetoplastid phylogenetics, with special reference to the evolution of parasitic trypanosomes. *Parasite* **15**, 226-232.

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

Stura, E. and Wilson, I. (1991). Applications of the streak seeding technique in protein crystallization. *Journal of Crystal Growth* **110**.

Svensjo, E., Cyrino, F. Z., Juliano, L. and Scharfstein, J. (1997). Plasma leakage induced in postcapillary venules by the major cysteine-proteinase from *Trypanosoma cruzi* and its modulation by H1-blocker mepyramine. *Microvasc Res* **54**, 93-97.

Symeonakis, E., Robinson, T. and Drake, N. (2007). GIS and multiple-criteria evaluation for the optimisation of tsetse fly eradication programmes. *Environ Monit Assess* **124**, 89-103.

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

Tastan Bishop, O. and Kroon, M. (2011). Study of protein complexes via homology modeling, applied to cysteine proteases and their protein inhibitors. *J Mol Model*.

Taylor, K. and Authié, E. (2004). Pathogenesis of Animal Trypanosomosis. In *The Trypanosomosis*: Oxfordshire: CABI Publishing.

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. *Exp Parasitol* **83**, 106-116.

Terry, R. J. (1957). Antibody against *Trypanosoma vivax* present in normal cotton rat serum. *Exp Parasitol* **6**, 404-411.

Tessier, P. M. and Lenhoff, A. M. (2003). Measurements of protein self-association as a guide to crystallization. *Curr Opin Biotechnol* **14**, 512-516.

Thekisoe, O. M., Kuboki, N., Nambota, A., Fujisaki, K., Sugimoto, C., Igarashi, I., Yasuda, J. and Inoue, N. (2007). Species-specific loop-mediated isothermal amplification (LAMP) for diagnosis of trypanosomosis. *Acta Trop* **102**, 182-189.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4354.

Tran, T., Buscher, 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. *J Biotechnol* **135**, 247-254.

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

Troeberg, L. (1997). Trypanopain: A possible target for anti-trypanosomal agents. In *Biochemistry*, PhD. Pietermaritzburg: University of KwaZulu-Natal.

Troeberg, L., Morty, R. E., Pike, R. N., Lonsdale-Eccles, J. D., Palmer, J. T., McKerrow, J. H. and Coetzer, T. H. (1999). Cysteine proteinase inhibitors kill cultured bloodstream forms of *Trypanosoma brucei brucei*. *Exp Parasitol* **91**, 349-355.

Turk, B., Turk, D. and Turk, V. (2000). Lysosomal cysteine proteases: more than scavengers. *Biochim Biophys Acta* **1477**, 98-111.

Turner, C. M. and Barry, J. D. (1989). High frequency of antigenic variation in *Trypanosoma brucei rhodesiense* infections. *Parasitology* **99 Pt 1**, 67-75.

Uzonna, J. E., Kaushik, R. S., Gordon, J. R. and Tabel, H. (1998). Experimental murine *Trypanosoma congolense* infections. I. Administration of anti-IFN-gamma antibodies alters trypanosome-susceptible mice to a resistant-like phenotype. *J Immunol* **161**, 5507-5015.

Vale, G. A. (1993). Development of baits for tsetse flies (Diptera: Glossinidae) in Zimbabwe. *J Med Entomol* **30**, 831-842.

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

Van den Bossche, P., de La Rocque, S., Hendrickx, G. and Bouyer, J. (2010). A changing environment and the epidemiology of tsetse-transmitted livestock trypanosomiasis. *Trends Parasitol* **26**, 236-243.

Varughese, K. I., Ahmed, F. R., Carey, P. R., Hasnain, S., Huber, C. P. and Storer, A. C. (1989). Crystal structure of a papain-E-64 complex. *Biochemistry* **28**, 1330-1332.

Vather, P. (2010). Vivapain: A cysteine peptidase from *Trypanosoma vivax*. MSc dissertation. University of KwaZulu-Natal (Pietermaritzburg campus).

Vernet, T., Berti, P. J., de Montigny, C., Musil, R., Tessier, D. C., Menard, R., Magny, M. C., Storer, A. C. and Thomas, D. Y. (1995). Processing of the papain precursor. The ionization state of a conserved amino acid motif within the Pro region participates in the regulation of intramolecular processing. *J Biol Chem* **270**, 10838-10846.

Vickerman, K. (1969). The fine structure of *Trypanosoma congolense* in its bloodstream phase. *J Protozool* **16**, 54-69.

Vickerman, K. (1978). Antigenic variation in trypanosomes. *Nature* **273**, 613-617.

Vincendeau, P. and Bouteille, B. (2006). Immunology and immunopathology of African trypanosomiasis. *An Acad Bras Cienc* **78**, 645-665.

Vreysen, M. J., Saleh, K. M., Ali, M. Y., Abdulla, A. M., Zhu, Z. R., Juma, K. G., Dyck, V. A., Msangi, A. R., Mkonyi, P. A. and Feldmann, H. U. (2000). *Glossina austeni* (*Diptera: Glossinidae*) eradicated on the island of Unguja, Zanzibar, using the sterile insect technique. *J Econ Entomol* **93**, 123-135.

Wang, B., Shi, G. P., Yao, P. M., Li, Z., Chapman, H. A. and Bromme, D. (1998). Human cathepsin F. Molecular cloning, functional expression, tissue localization, and enzymatic characterization. *J Biol Chem* **273**, 32000-32008.

Wint, W. and Rogers, D. (2000). *Predicted Distributions of Tsetse in Africa*. Animal Health Service of the Animal Production and Health Division of the Food and Agriculture Organization of the United Nations., Rome, 2011.

Wiser, M. F., Lonsdale-Eccles, J. D., D'Alessandro, A. and Grab, D. J. (1997). A cryptic protease activity from Trypanosoma cruzi revealed by preincubation with kininogen at low temperatures. *Biochem Biophys Res Commun* **240**, 540-544.

Woo, P. T. (1970). The haematocrit centrifuge technique for the diagnosis of African trypanosomiasis. *Acta Trop* **27**, 384-386.

Wu, S. and Letchworth, G. J. (2004). High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol. *Biotechniques* **36**, 152-154.

Yadav, S. C., Kumar, R., Kumar, S., Tatu, U., Singh, R. K. and Gupta, A. K. (2011). Identification and characterization of cysteine proteinases of Trypanosoma evansi. *Parasitol Res*.

Yongsheng, Y., Yongchun, O., Chengmai, R., Yuanguo, C. and Fenqin, Z. (1996). Trypanocidal value of liposomal diminazene in experimental *Trypanosoma brucei evansi* infection in mice. *Vet Parasitol* **61**, 349-352.

Zeelen, J. P. (2009). Interpretation of the crystallisation drop results. In *Protein Crystallisation, 2nd ed.* La Jolla: International University Line.

APPENDIX 1: Screening conditions used in JSCG+ and PACT commercial screens

JSCG+ screen:

Tube No.	Salt	Buffer	pH	Precipitant
1	0.2 M lithium sulfate	0.1 M sodium acetate	4.5	50 % v/v PEG 400
2	None	0.1 M sodium citrate	5.5	20 % w/v PEG 3000
3	0.2 M di-ammonium hydrogen citrate	None	-	20 % w/v PEG 3350
4	0.02 M calcium chloride	0.1 M sodium acetate	4.6	30 % v/v MPD
5	0.2 M magnesium formate	None	-	20 % w/v PEG 3350
6	0.2 M lithium sulfate	0.1 M phosphate/citrate	4.2	20 % w/v PEG 1000
7	None	0.1 M CHES	9.5	20 % w/v PEG 8000
8	0.2 M ammonium formate	None	-	20 % w/v PEG 3350
9	0.2 M ammonium chloride	None	-	20 % w/v PEG 3350
10	0.2 M potassium formate	None	-	20 % w/v PEG 3350
11	0.2 M ammonium dihydrogen phosphate	0.1 M Tris	8.5	50 % v/v MPD
12	0.2 M potassium nitrate	None	-	20 % w/v PEG 3350
13	None	0.1 M citrate	4.0	0.8 M ammonium sulfate
14	0.2 M sodium thiocyanate	None	-	20 % w/v PEG 3350
15	None	0.1 M Bicine	9.0	20 % w/v PEG 6000
16	None	0.1 M HEPES	7.5	10 % w/v PEG 8000/ 8 % v/v Ethylene glycol
17	None	0.1 M sodium cacodylate	6.5	40 % v/v MPD/ 5 % w/v PEG 8000
18	None	0.1 M phosphate/citrate	4.2	40 % v/v Ethanol/ 5 % w/v PEG 1000
19	None	0.1 M sodium acetate	4.6	8 % w/v PEG 4000
20	0.2 M magnesium chloride	0.1 M Tris	7.0	10 % w/v PEG 8000
21	None	0.1 M citrate	5.0	20 % w/v PEG 6000
22	0.2 M magnesium chloride	0.1 M sodium cacodylate	6.5	50 % v/v PEG 200
23	None	None	6.5	1.6 M tri-sodium citrate
24	0.2 M tri-potassium citrate	None	-	20 % w/v PEG 3350
25	0.2 M sodium chloride	0.1 M phosphate/citrate	4.2	20 % w/v PEG 8000
26	1.0 M lithium chloride	0.1 M Na citrate	4.0	20 % w/v PEG 6000
27	0.2 M ammonium nitrate	None	-	20 % w/v PEG 3350
28	None	0.1 M Na HEPES	7.0	10 % w/v PEG 6000
29	None	0.1 M Na HEPES	7.5	0.8 M sodium dihydrogen phosphate 0.8 M potassium dihydrogen phosphate
30	None	0.1 M phosphate/citrate	4.2	40 % v/v PEG 300
31	0.2 M zinc acetate	0.1 M sodium acetate	4.5	10 % w/v PEG 3000
32	None	0.1 M Tris	8.5	20 % v/v Ethanol
33	None	0.1 M Na/K phosphate	6.2	25 % v/v 1,2-propanediol 10 % v/v Glycerol
34	None	0.1 M Bicine	9.0	10 % w/v PEG 20,000/ 2% v/v Dioxane
35	None	0.1 M sodium acetate	4.6	2.0 M ammonium sulfate
36	None	None	-	10 % w/v PEG 1000/ 10 % w/v PEG 8000
37	None	None	-	24 % w/v PEG 1500/ 20 % v/v Glycerol
38	0.2 M magnesium chloride	0.1 M Na HEPES	7.5	30 % v/v PEG 400
39	0.2 M sodium chloride	0.1 M Na/K phosphate	6.2	50 % v/v PEG 200
40	0.2 M lithium sulfate	0.1 M sodium acetate	4.5	30 % w/v PEG 8000
41	None	0.1 M HEPES	7.5	70 % v/v MPD
42	0.2 M magnesium chloride	0.1 M Tris	8.5	20 % w/v PEG 8000
43	0.2 M lithium sulfate	0.1 M Tris	8.5	40 % v/v PEG 400
44	None	0.1 M Tris	8.0	40 % v/v MPD
45	0.17 M ammonium	None	-	25.5 % w/v PEG 4000/

	sulfate			15 % v/v Glycerol
46	0.2 M calcium acetate	0.1 M sodium cacodylate	6.5	40 % v/v PEG 300
47	0.14 M calcium chloride	0.07 M sodium acetate	4.6	14 % v/v 2-propanol/ 30 % v/v Glycerol
48	0.04 M potassium dihydrogen phosphate	None	-	16 % w/v PEG 8000/ 20 % v/v Glycerol
49	None	0.1 M sodium cacodylate	6.5	1.0 M tri-sodium citrate
50	0.2 M sodium chloride	0.1 M sodium cacodylate	6.5	2.0 M ammonium sulfate
51	0.2 M sodium chloride	0.1 M HEPES	7.5	10 % v/v 2-propanol
52	0.2 M lithium sulfate	0.1 M Tris	8.5	1.26 M ammonium sulfate
53	None	0.1 M CAPS	10.5	40 % v/v MPD
54	0.2 M zinc acetate	0.1 M imidazole	8.0	20 % w/v PEG 3000
55	0.2 M zinc acetate	0.1 M sodium cacodylate	6.5	10 % v/v 2-propanol
56	None	0.1 M sodium acetate	4.5	1.0 M di-ammonium hydrogen phosphate
57	None	0.1 M MES	6.5	1.6 M magnesium sulfate
58	None	0.1 M Bicine	9.0	10 % w/v PEG 6000
59	0.16 M calcium acetate	0.08 M sodium cacodylate	6.5	14.4 % w/v PEG 8000/ 20 % v/v glycerol
60	None	0.1 M imidazole	8.0	10 % w/v PEG 8000
61	0.05 M caesium chloride	0.1 M MES	6.5	30 % v/v Jeffamine M-600
62	None	0.1 M Na Citrate	5.0	3.2 M ammonium sulfate
63	None	0.1 M Tris	8.0	20 % v/v MPD
64	None	0.1 M HEPES	7.5	20 % v/v Jeffamine M-600
65	0.2 M magnesium chloride	0.1 M Tris	8.5	50 % v/v ethylene glycol
66	None	0.1 M Bicine	9.0	10 % v/v MPD
67	None	None	7.0	0.8 M succinic acid
68	None	None	7.0	2.1 M DL-malic acid
69	None	None	7.0	2.4 M sodium malonate
70	1.1 M sodium malonate	0.1 M HEPES	7.0	0.5 % v/v Jeffamine ED-2001
71	1.0 M succinic acid	0.1 M HEPES	7.0	1 % w/v PEG 2000 MME
72	None	0.1 M HEPES	7.0	30 % v/v Jeffamine M-600
73	None	0.1 M HEPES	7.0	30 % v/v Jeffamine ED-2001
74	0.02 M magnesium chloride	0.1 M HEPES	7.5	22 % w/v polyacrylic acid 5100 sodium salt
75	0.01 M cobalt chloride	0.1 M Tris	8.5	20 % w/v polyvinylpyrrolidone K15
76	0.2 M tri-methylamine N-oxide	0.1 M Tris	8.5	20 % w/v PEG 2000 MME
77	0.005 M cobalt chloride 0.005 M cadmium chloride 0.005 M magnesium chloride 0.005 M nickel chloride	0.1 M HEPES	7.5	12 % w/v PEG 3350
78	0.2 M sodium malonate	None	7.0	20 % w/v PEG 3350
79	0.1 M succinic acid	None	7.0	15 % w/v PEG 3350
80	0.15 M DL - malic acid	None	7.0	20 % w/v PEG 3350
81	0.1 M potassium thiocyanate	None	-	30 % w/v PEG 2000 MME
82	0.15 M potassium bromide	None	-	30 % w/v PEG 2000 MME
83	None	0.1 M Bis Tris	5.5	2.0 M ammonium sulfate
84	None	0.1 M Bis Tris	5.5	3.0 M sodium chloride
85	None	0.1 M Bis Tris	5.5	0.3 M magnesium formate
86	1.0 M ammonium sulfate	0.1 M Bis Tris	5.5	1 % w/v PEG 3350
87	None	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
88	0.2 M calcium chloride	0.1 M Bis Tris	5.5	45 % v/v MPD
89	0.2 M ammonium acetate	0.1 M Bis Tris	5.5	45 % v/v MPD
90	0.1 M ammonium acetate	0.1 M Bis Tris	5.5	17 % w/v PEG 10000
91	0.2 M ammonium sulfate	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
92	0.2 M sodium chloride	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
93	0.2 M lithium sulfate	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
94	0.2 M ammonium acetate	0.1 M Bis Tris	5.5	25 % w/v PEG 3350

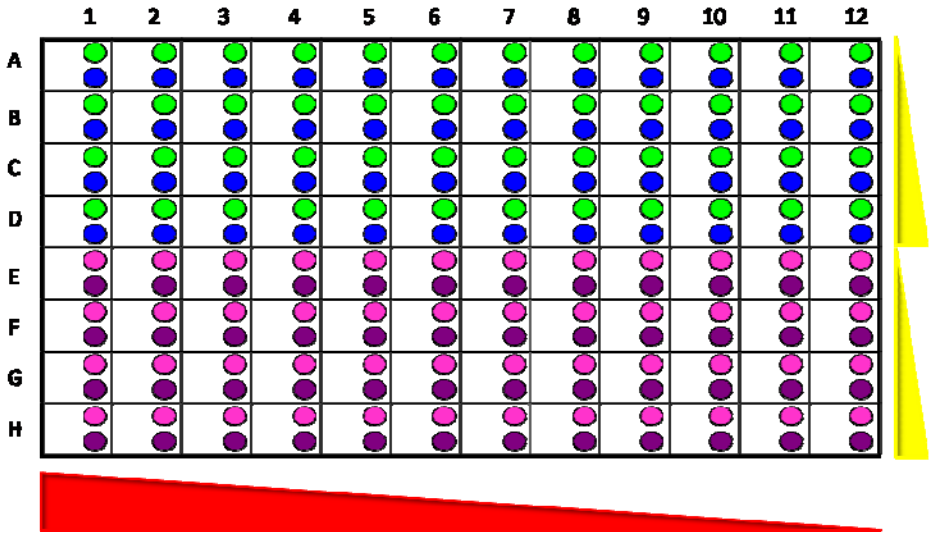
95	0.2 M magnesium chloride	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
96	0.2 M ammonium acetate	0.1 M HEPES	7.5	45 % v/v MPD

PACT screen:

Tube No.	Salt	Buffer	pH	Precipitant
1		0.1 M SPG buffer	4	25% (w/v) PEG 1500
2		0.1 M SPG buffer	5	25% (w/v) PEG 1500
3		0.1 M SPG buffer	6	25% (w/v) PEG 1500
4		0.1 M SPG buffer	7	25% (w/v) PEG 1500
5		0.1 M SPG buffer	8	25% (w/v) PEG 1500
6		0.1 M SPG buffer	9	25% (w/v) PEG 1500
7	0.2 M Sodium chloride	0.1 M Sodium acetate	5	20% (w/v) PEG 6000
8	0.2 M Ammonium chloride	0.1 M Sodium acetate	5	20% (w/v) PEG 6000
9	0.2 M Lithium chloride	0.1 M Sodium acetate	5	20% (w/v) PEG 6000
10	0.2 M Magnesium chloride	0.1 M Sodium acetate	5	20% (w/v) PEG 6000
11	0.2 M Calcium chloride	0.1 M Sodium acetate	5	20% (w/v) PEG 6000
12	0.01 M Zinc chloride	0.1 M Sodium acetate	5	20% (w/v) PEG 6000
13		0.1 M MIB buffer	4	25% (w/v) PEG 1500
14		0.1 M MIB buffer	5	25% (w/v) PEG 1500
15		0.1 M MIB buffer	6	25% (w/v) PEG 1500
16		0.1 M MIB buffer	7	25% (w/v) PEG 1500
17		0.1 M MIB buffer	8	25% (w/v) PEG 1500
18		0.1 M MIB buffer	9	25% (w/v) PEG 1500
19	0.2 M Sodium chloride	0.1 M MES	6	20% (w/v) PEG 6000
20	0.2 M Ammonium chloride	0.1 M MES	6	20% (w/v) PEG 6000
21	0.2 M Lithium chloride	0.1 M MES	6	20% (w/v) PEG 6000
22	0.2 M Magnesium chloride	0.1 M MES	6	20% (w/v) PEG 6000
23	0.2 M Calcium chloride	0.1 M MES	6	20% (w/v) PEG 6000
24	0.01 M Zinc chloride	0.1 M MES	6	20% (w/v) PEG 6000
25		0.1 M PCB buffer	4	25% (w/v) PEG 1500
26		0.1 M PCB buffer	5	25% (w/v) PEG 1500
27		0.1 M PCB buffer	6	25% (w/v) PEG 1500
28		0.1 M PCB buffer	7	25% (w/v) PEG 1500
29		0.1 M PCB buffer	8	25% (w/v) PEG 1500
30		0.1 M PCB buffer	9	25% (w/v) PEG 1500
31	0.2 M Sodium chloride	0.1 M HEPES	7	20% (w/v) PEG 6000
32	0.2 M Ammonium chloride	0.1 M HEPES	7	20% (w/v) PEG 6000
33	0.2 M Lithium chloride	0.1 M HEPES	7	20% (w/v) PEG 6000
34	0.2 M Magnesium chloride	0.1 M HEPES	7	20% (w/v) PEG 6000
35	0.2 M Calcium chloride	0.1 M HEPES	7	20% (w/v) PEG 6000
36	0.01 M Zinc chloride	0.1 M HEPES	7	20% (w/v) PEG 6000
37		0.1 M MMT buffer	4	25% (w/v) PEG 1500
38		0.1 M MMT buffer	5	25% (w/v) PEG 1500
39		0.1 M MMT buffer	6	25% (w/v) PEG 1500
40		0.1 M MMT buffer	7	25% (w/v) PEG 1500
41		0.1 M MMT buffer	8	25% (w/v) PEG 1500
42		0.1 M MMT buffer	9	25% (w/v) PEG 1500
43	0.2 M Sodium chloride	0.1 M Tris	8	20% (w/v) PEG 6000
44	0.2 M Ammonium chloride	0.1 M Tris	8	20% (w/v) PEG 6000
45	0.2 M Lithium chloride	0.1 M Tris	8	20% (w/v) PEG 6000
46	0.2 M Magnesium chloride	0.1 M Tris	8	20% (w/v) PEG 6000
47	0.2 M Calcium chloride	0.1 M Tris	8	20% (w/v) PEG 6000
48		0.1 M Tris	8	20% (w/v) PEG 6000
49	0.2 M Sodium fluoride			20% (w/v) PEG 3350
50	0.2 M Sodium bromide			20% (w/v) PEG 3350
51	0.2 M Sodium iodide			20% (w/v) PEG 3350
52	0.2 M Potassium thiocyanate			20% (w/v) PEG 3350
53	0.2 M Sodium nitrate			20% (w/v) PEG 3350
54	0.2 M Sodium formate			20% (w/v) PEG 3350
55	0.2 M Sodium acetate			20% (w/v) PEG 3350
56	0.2 M Sodium sulfate			20% (w/v) PEG 3350
57	0.2 M Potassium/sodium tartrate			20% (w/v) PEG 3350
58	0.2 M Sodium/potassium phosphate			20% (w/v) PEG 3350
59	0.2 M Sodium citrate			20% (w/v) PEG 3350
60		0.2 M Sodium malonate		20% (w/v) PEG 3350
61	0.2 M Sodium fluoride	0.1 M Bis Tris propane	6.5	20% (w/v) PEG 3350
62	0.2 M Sodium bromide	0.1 M Bis Tris propane	6.5	20% (w/v) PEG 3350
63	0.2 M Sodium iodide	0.1 M Bis Tris propane	6.5	20% (w/v) PEG 3350

64	0.2 M Potassium thiocyanate	0.1 M Bis Tris propane	6.5		20% (w/v) PEG 3350
65	0.2 M Sodium nitrate	0.1 M Bis Tris propane	6.5		20% (w/v) PEG 3350
66	0.2 M Sodium formate	0.1 M Bis Tris propane	6.5		20% (w/v) PEG 3350
67	0.2 M Sodium acetate	0.1 M Bis Tris propane	6.5		20% (w/v) PEG 3350
68	0.2 M Sodium sulfate	0.1 M Bis Tris propane	6.5		20% (w/v) PEG 3350
69	0.2 M Potassium/sodium tartrate	0.1 M Bis Tris propane	6.5		20% (w/v) PEG 3350
70	0.2 M Sodium/potassium phosphate	0.1 M Bis Tris propane	6.5		20% (w/v) PEG 3350
71	0.2 M Sodium citrate	0.1 M Bis Tris propane	6.5		20% (w/v) PEG 3350
72	0.2 M Sodium malonate	0.1 M Bis Tris propane	6.5		20% (w/v) PEG 3350
73	0.2 M Sodium fluoride	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
74	0.2 M Sodium bromide	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
75	0.2 M Sodium iodide	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
76	0.2 M Potassium thiocyanate	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
77	0.2 M Sodium nitrate	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
78	0.2 M Sodium formate	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
79	0.2 M Sodium acetate	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
80	0.2 M Sodium sulfate	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
81	0.2 M Potassium/sodium tartrate	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
82	0.2 M Sodium/potassium phosphate	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
83	0.2 M Sodium citrate	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
84	0.2 M Sodium malonate	0.1 M Bis Tris propane	7.5		20% (w/v) PEG 3350
85	0.2 M Sodium fluoride	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
86	0.2 M Sodium bromide	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
87	0.2 M Sodium iodide	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
88	0.2 M Potassium thiocyanate	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
89	0.2 M Sodium nitrate	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
90	0.2 M Sodium formate	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
91	0.2 M Sodium acetate	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
92	0.2 M Sodium sulfate	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
93	0.2 M Potassium/sodium tartrate	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
94	0.2 M Sodium/potassium phosphate	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
95	0.2 M Sodium citrate	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350
96	0.2 M sodium malonate	0.1 M Bis Tris propane	8.5		20% (w/v) PEG 3350

APPENDIX 2: Optimisation screen design for *Tco*CATL (K39F; E44P).



KEY:

PEG 3350/PEG 6000 (36%-15%)

pH of 0.1 M Sodium citrate buffer (pH 5.0-5.5 for A-D and pH 5.0-5.5 for E-H)

Protein:crystallant in subwells:

- 80:20 subwell 1 (size of drop is 0.6 µl)
- 47:20 subwell 2 (size of drop is 0.6 µl)
- 40:20 subwell 1 (size of drop is 0.6 µl)
- 30:20 subwell 2 (size of drop is 0.6 µl)

The top half of the plate (A-D) had protein:crystallant ratio of 80:20 and 47:20. The second half of the plate (E-H) had 40:20 and 30:20 ratios. The range of buffer pH was also split for each half of the plate (increments of 0.17 in pH value) . The PEG concentration decreased over the 12 columns. Two plates were prepared, one for PEG 3350 and the other for PEG 6000.