Characterization of CAP5.5 and CAP5.5V, atypical calpains in *Trypanosoma brucei*

Tenzing Tashi Sherpa Lama

Saint Edmund Hall

A thesis submitted for the degree of Master of Science (by Research) in the University of Oxford, Hilary Term 2013.

Acknowledgements

I would like to thank Keith Gull for his advice, mentorship and guidance. I would also like to thank him for the opportunity to work in his laboratory.

Samuel Dean, for his generosity in investing his time in my project and me; his patience in dealing with my many questions and for his contributions towards characterizing my proteins (Fig. 4.8 and 4.9B- 4.9D).

Pietro Roversi, for his helpful suggestions and for lending his expertise particularly in the large-scale purification of proteins. Many thanks also to Steven Johnson for his help with the MALLS data.

I would also like to thank all members of the Gull lab for their camaraderie and invaluable advice. I am grateful to Susan Lea and members of her lab for accommodating me and allowing me to use their equipments. Many thanks also to Peter Davies (Queen's University, Canada) for his gift of bacterial stabilates expressing mouse CAPN2.

Lastly, I would like to thank the Keasbey Memorial Foundation for funding my study at Oxford.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or institute of learning.

Abstract

Calpains are a group of calcium-dependent cysteine proteases with roles ranging from cytoskeletal remodelling to signal transduction. In *Trypanosoma brucei*, the causal agent of African sleeping sickness, an unusually large number of calpain-like proteins exist, whose functions remain unknown. Recently, the procyclic form-specific cytoskeleton associated protein 5.5 (CAP5.5) and its bloodstream-specific paralogous variant, CAP5.5V were discovered to be essential for the proper morphogenesis of the parasite. However, precise functional roles of these proteins remain unknown. In this MSc thesis, I employed a two-pronged approach combining cell biology studies of CAP5.5 and CAP5.5V in *T. brucei* with *in vitro* assays of recombinant calpains expressed in *Escherichia coli* to further characterize CAP5.5 and CAP5.5V. To resolve whether the function of CAP5.5 and CAP5.5V is dependent upon proteolytic activity, I expressed and purified recombinant CAP5.5V's catalytic domain. I demonstrated that this domain did not have any detectable autolytic activity or activity against universal protease substrates such as casein. Ectopic expression of CAP5.5::YFP::Ty in procyclic trypanosomes led to nuclear mis-positioning among other morphological defects. Additionally, I demonstrated a cell cycle dependency on incorporation of CAP5.5::YFP::Ty into the cytoskeleton and corroborate previous data that suggest that CAP5.5 is built in from the dynamic, posterior end of the cell. Investigation of the molecular determinants in CAP5.5's sequence such as motifs or domains suggested that CAP5.5's domain IV may be the domain mediating association with the sub-pellicular cytoskeleton.

Abbreviations

ATP	Adenosine triphosphate
BSF	Bloodstream form
BLAST	Basic local alignment search tool
BME	ß-mercaptoethanol
bp	Base pair
CAP	Cytoskeleton associated protein
CAPN	Calpain
DABCO	1,4-diazabicyclo [2.2.2] octane
DAPI	4', 6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
EDTA	Ethylenedinitrilo tetraacetic acid (disodium salt)
EGTA	Ethylenebis (oxyethylenenitrilo) tetraacetic acid
ER	Endoplasmic reticulum
FAZ	Flagellum attachment zone
FCS	Fetal calf serum
GFP	Green fluorescent protein
GPI	Glycosylphosphatidylinositol
HAT	Human African trypanosomiasis
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K	Kinetoplast
kb	Kilobase
kDa	Kilo-dalton
kDNA	Kinetoplast DNA
LB	Luria broth
MALLS	Multi-angle laser light scattering
MAP	Microtubule associated protein
N	Nucleus
ORF	Open reading frame
PBS	Phosphate buffer saline

PCF	Procyclic form
PCR	Polymerase chain reaction
PFC	Paraflagellar rod proteome component
PFR	Paraflagellar rod
PIPES	Piperazine-N, N-bis (2-ethane sulphonic acid)
rRNA	Ribosomal RNA
RNA	Ribonucleic acid
RNAi	RNA interference
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SMP	Small myristoylated protein
SKCRP	Small kinetoplastid calpain-related protein
Tet	Tetracycline
VSG	Variant surface glycoprotein
WCB	Whole cell body
WHO	World Health Organization

Table of Contents

ACKNOWLEDGEMENTS ..... **II**
DECLARATION ..... **III**
ABSTRACT ..... **IV**
ABBREVIATIONS ..... **V**
TABLE OF CONTENTS ..... **VII**
FIGURES AND TABLES ..... **X**

1. INTRODUCTION **1**
1.1 Trypanosoma brucei *2*
1.1.1 Trypanosomiasis *2*
1.1.2 Trypanocidal drugs *3*
1.1.3 Life cycle of T. brucei *4*
1.1.4 Cell cycle of procyclic T. brucei *8*
1.1.5 Cytoskeleton *11*
1.1.5.1 Microtubule-associated proteins *12*
1.1.6 Gene Expression *15*
1.2 Calpains *16*
1.2.1 Typical calpains – structure *17*
1.2.2 Typical calpains – function *21*
1.2.3 Atypical calpains *23*
1.2.4 Kinetoplastid calpains *26*
1.2.5 Localization and biochemical characteristics of kinetoplastid calpains *30*
1.2.6 Acylation *32*
1.2.7 CAP5.5 and CAP5.5V *34*
1.2.8 Aims and strategies *37*

2. MATERIALS AND METHODS **38**
2.1 Molecular Biology *39*
2.1.1 Polymerase Chain Reaction *39*
2.1.2 Agarose gel electrophoresis *39*
Restriction sites are in italics and in a 5' to 3' format. All primers were ordered from Sigma. *40*
2.1.3 Restriction digest and DNA ligation *41*
2.1.4 DNA isolation *42*
2.2 Heterologous protein expression and purification *42*
2.2.1 Bacterial expression vectors and host strain *42*
2.2.2 Protein expression conditions *45*
2.2.3 Nickel affinity chromatography *45*
2.2.4 Size exclusion chromatography and MALLS *46*

2.2.5 Crystallization screening *47*
2.2.6 Protease activity assays *47*
2.2.7 Polyacrylamide gel electrophoresis *48*
2.2.8 Optimization trials *49*
2.3 T. brucei cell lines and cell culture *50*
2.3.1 Expression vectors *50*
2.3.2 Trypanosome cell culture *52*
2.3.3 Transfection of trypanosomes *53*
2.4 Protein biochemistry (T. brucei) *54*
2.4.1 Preparation of protein samples *54*
2.4.2 Western blot and immunodetection *54*
2.5 Light Microscopy *55*
2.5.1 Native fluorescence microscopy *56*
2.5.2 Immunofluorescence microscopy *56*
2.6 Bioinformatics *57*
2.6.1 Amino acid alignment *57*
2.6.2 Homology models *57*

3. EXPRESSION, PURIFICATION AND BIOCHEMICAL ANALYSES OF RECOMBINANT CALPAINS 59
3.1 Introduction *60*
3.2 Bioinformatic analyses of T. brucei CAP5.5 and CAP5.5V *60*
3.3 Soluble CAP5.5 and CAP5.5V expression *66*
3.4 Heterodisperse aggregation of CAP5.5 and CAP5.5V *71*
3.5 Optimization of protein purification to resolve aggregation *75*
3.6 Expression and purification of CAP5.5V domain II *84*
3.7 Crystallization trials of CAP5.5V domain II *87*
3.8 CAP5.5V domain II has no detectable protease activity *88*
3.9 Summary *91*

4. CHARACTERIZATION OF CAP5.5 AND CAP5.5V IN *T. BRUCEI* 93
4.1 Introduction *94*
4.2 Expression constructs *95*
4.3 Over-expression of CAP5.5 in procyclic T. brucei *96*
4.4 Temporal expression and localization of CAP5.5 ..... *100*
4.5 Characterization of CAP5.5 N6::YFP::Ty *105*
4.6 Characterization of CAP5.5 N37::YFP::Ty *110*
4.7 Characterization of CAP5.5 domain I::YFP::Ty *112*
4.8 Localization of domain I-II::YFP::Ty *113*
4.9 Expression and localization of domain I-III::YFP::Ty *116*
4.10 Characterization of GFP::CAP5.5 and GFP::CAP5.5V *119*
4.11 Over-expression of CAP5.5 in procyclic cells *128*
4.12 Summary *137*

5. DISCUSSION 140
5.1 Introduction.....141
5.2 Aggregation of recombinant calpains.....143
5.3 CAP5.5V D-II has no detectable protease activity.....145
5.4 Localization and targeting of CAP5.5 148
5.5 Posterior-end incorporation of CAP5.5 into the procyclic cytoskeleton.....153
5.6 Cell cycle-regulated incorporation of CAP5.5 into the procyclic cytoskeleton 155
5.7 Functional characterization of CAP5.5.....156
5.8 Conclusion 159

6. REFERENCES.....161

Figures and Tables

Chapter 1. Introduction

Figure 1.1. Overview of the life cycle of Trypanosoma brucei.....6
Figure 1.2. Cell cycle of procyclic T. brucei. 10
Figure 1.3. Domain architecture, nomenclature and structure of typical calpain, CAPN2. 19
Figure 1.4. Classification and domain architecture of kinetoplastids calpains. 29

Table 1.1. Biochemical features and localization of the hitherto characterized kinetoplastid calpains. 31

Chapter 2. Materials and Methods

Figure 2.1. Vector map of pET29a used in this thesis. 43
Figure 2.2. Vector map of pET24a 51

Table 2.1 List of primers used in this thesis work. 40

Chapter 3. Expression, purification and biochemical analyses of recombinant calpains

Figure 3.1. Amino acid alignment of the catalytic domains of human CAPN1 and CAPN2 with the corresponding domains of CAP5.5 and CAP5.5V.. 62
Figure 3.2. Homology modelling and domain parsing of CAP5.5V.....65
Figure 3.3. Schematic representation of the expression constructs made in this study. 67
Figure 3.4. SDS-PAGE and protease assay profiles of CAP5.5 and CAP5.5V……....70
Figure 3.5 Aggregation of CAP5.5V.....73
Figure 3.6. Molecular mass determination of CAP5.5V size exclusion chromatography eluates.....76
Figure 3.7. Assaying for CAP5.5V aggregation under varying pH and reducing conditions. 79
Figure 3.8. Size exclusion chromatography elution profile of recombinant CAP5.5V following iodoacetamide treatment.....80
Figure 3.8. SDS-PAGE analyses of protein expression. 82
Figure 3.9. Soluble expression of CAP5.5V domain II visualized by SDS-PAGE.....83
Figure 3.10. Elution profile and corresponding SDS-PAGE analyses of the fractions. 85
Figure 3.11. Molecular mass determination of CAP5.5V domain II under two conditions of varying ionic strength. 86

Figure 3.12. CAP5.5V domain II has no protease activity......89

Table 3.1. Biochemical properties of CAP5.5 and CAP5.5V. 63

Chapter 4. Characterization of CAP5.5 and CAP5.5V in *T. brucei*

Figure 4.1. Schematic representation of the expression constructs used in this chapter. 97
Figure 4.2. Over-expression of CAP5.5::YFP::Ty in procyclic cells......99
Figure 4.3. Temporal localization of CAP5.5::YFP::Ty following induction......103
Figure 4.4. N6::YFP::Ty localizes to the cell body and flagellum......108
Figure 4.5. N37::YFP::Ty is restricted to the cell body......111
Figure 4.6. Ectopic expression of DI::YFP::Ty in procyclic cells. ……..114
Figure 4.7. Ectopic expression of CAP5.5 domains I-II......115
Figure 4.8. Ectopic expression of domains I-III results in a growth defect phenotype.. 118
Figure 4.9. Ectopic expression of GFP::CAP5.5 and GFP::CAP5.5V lead to growth defects. 123
Figure 4.10. GFP::CAP5.5V localizes to the cytoskeleton of procyclic cells......124
Figure 4.11. Over-expression of CAP5.5 in procyclic parasites leads to abnormal cell types. 127
Figure 4.12. Morphometric analyses of procyclic cells over-expressing CAP5.5. ...130
Figure 4.13. Morphometric analyses of 1K1N procyclic cells over-expressing CAP5.5.......133
Figure 4.14. A summary of the localization of the ectopic expression constructs used in this study. 138

Table 4.1. Morphometric measurements of 1K1N cells over-expressing CAP5.5 48 hours post induction......134

1. Introduction

1.1 *Trypanosoma brucei*

1.1.1 Trypanosomiasis

Although descriptions of trypanosomiasis have been found even in the Petri papyri dating from the second millennium BC, it was only in 1895 that David Bruce implicated the parasite, *Trypanosoma brucei,* as the etiological agent of this crippling disease (Griffith, 1898; Bruce, 1895; Steverding, 2008). *T. brucei* belongs to the phylum *Euglenozoa*, the class *Kinetoplastida*, the order *Trypanosomatidae* and the genus *Trypanosoma. T. brucei* spp (comprising *T. brucei brucei* and the human infective forms *T. brucei gambiense* and *T. brucei rhodesiense*) are responsible for the human trypanosomiasis (HAT) also called African sleeping sickness, which has a devastating effect on the physical and economic health of the people in sub-Saharan Africa. In 2010, 7139 cases of infection with the protozoan parasite were reported after years of efforts to control this disease (WHO). The parasite also infects cattle and game animals, where the disease is called 'nagana', further contributing to the endemic poverty of the region. Of the four drugs currently used to treat sleeping sickness, most of them are toxic and ineffective at eradicating parasitic infection (reviewed in Docampo & Moreno, 2003). This disease, often characterized by bouts of fever, joint pains and neurological disturbances eventually leads to coma and death if left untreated (Forde, 1904). There is little hope of a vaccine for African sleeping sickness due to antigenic variation and a desperate need for modern drug therapies against this neglected tropical disease.

My thesis focuses on calpains, calcium-regulated cysteine proteases, in *T. brucei* that are essential for the proper morphogenesis of the parasite. In this

introduction, I shall lay the context by providing a global overview of the parasite's life and cell cycles, its sub-pellicular cytoskeleton and gene expression, followed by a review of the current literature on both typical and atypical calpains, especially those found in kinetoplastid parasites.

1.1.2 Trypanocidal drugs

The drugs against the two human-infective *T. brucei* hemoflagellates, *T. b. rhodesiense* (East African or Rhodesian form) and *T. b. gambiense* (West African or Gambian form) are administered according to the clinical stages of the disease. In the early or hemolymphatic stage, the parasites are found in the bloodstream and lymphatic system following the bite of an infected tsetse fly. Pathologically, this stage is characterized by bouts of fever, headaches, arthralgia, and fatigue (Pentreath, 1995). Without treatment during this early stage, the parasites can cross the blood-brain barrier and invade the central nervous system in the advanced or the encephalitic stage. The range of neurological and psychiatric symptoms experienced at this stage includes sleep and motor disturbances, meningoencephalitis, and severe headaches, often culminating in coma and death (Bacchi, 2009).

Two of the main early-stage drugs against *T. b. gambiense* and *T. b. rhodesiense* are pentamidine and suramin respectively (reviewed in Kennedy, 2006). Suramin can inhibit several enzymes including those in the glycolytic and pentose phosphate pathways whereas pentamidine interferes with the parasite's nucleic acid metabolism (Barrett *et al.,* 2007; Wang, 1995). While these two drugs can generally be effective, they do elicit several undesirable effects such as hypotension and

hyper/hypoglycemia, renal failure, anaphylactic shock, and neurological complications among others (reviewed in Kennedy, 2013).

The drugs for the advanced stage are more toxic and complicated to administer. Melarsoprol, an arsenic derivative in propylene glycol, is used primarily for late-stage *T. b. rhodesiense* infection although it is also effective against late-stage *T. b. gambiense* infection. Toxicity remains a huge concern with its administration, as 10% of treated patients develop encephalopathies, pulmonary edema and oftentimes, die (Gehrig & Efferth, 2008). The mechanism of melarsoprol action is not very clear even though it has been actively studied. The other drug for this stage is eflorninthine, an ornithine decarboxylase inhibitor, which is only effective against *T.b. gambiense* (Kennedy, 2004; Kennedy, 2008). While effective, an eflorninthine regimen is often strict and complicated to follow (Bacchi, 2009). For an extensive review of the history of these drugs including those currently in clinical trials, please refer to Kennedy (2013).

1.1.3 Life cycle of *T. brucei*

The life cycle of *T. brucei* takes place in the mammalian host bloodstream and inside the insect vector, tsetse fly -*Glossina* spp.- accompanied by great changes in morphology, metabolism, gene expression and signalling pathways in an adaptive response of these single-celled eukaryotes to the various environments they encounter (Fig. 1.1).

In the mammalian bloodstream, *T. brucei* can be found extracellularly as proliferative and morphologically long-slender trypomastigotes with their kinetoplasts at the terminal posterior ends (Robertson, 1913). These cells rely primarily on the energy from the glycolytic pathway converting glucose from the host environment to pyruvate in specialized organelles called glycosomes. Its mitochondrial ability for oxidative phosphorylation is repressed except for the plant-like alternative oxidase, which maintains the glycosomal redox balance (Bringaud *et al.,* 2006).

Against the backdrop of a lytic immunoglobin response from the mammalian host, the parasites are able to evade an immune response by a controlled monoallelic expression of a dense layer of variant surface glycoprotein (VSG) linked to the surface membrane by a glycosylphosphatidylinositol (GPI) anchor (Borst, 2002). This VSG coating also protects against complement-mediated lysis and shields invariant surface glycoproteins from host antibodies by steric hindrance (Ferrante & Allison, 1983; Ziegelbauer & Overath, 1993). Antigenic variation coincides with periodic waves of parasitemia seen in infected patients, where the parasite establishes an antigenically distinct population following the humoral response against parasites expressing the previous VSG (McCulloch, 2004; Pays *et al.,* 2004).

As the infection progresses and parasitemia increases, the slender bloodstream forms differentiate into morphologically short-stumpy, quiescent trypomastigotes via quorum-sensing in response to a signalling factor called stumpy-induction factor (Vassella *et al.,* 1997; Reuner *et al.,* 1997). These stumpy forms are uniformly arrested in the G1/G0 phase, limiting further increase in parasitemia, thereby

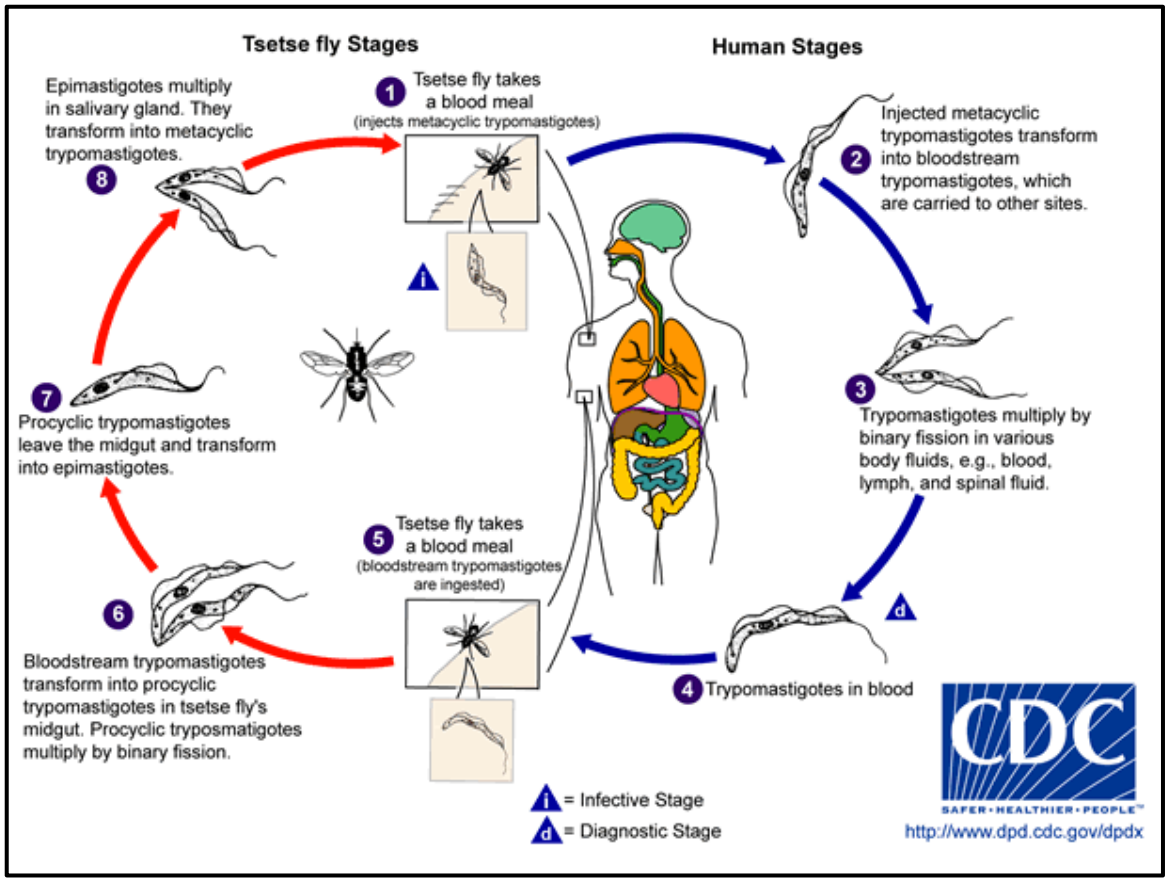


Figure 1.1. Overview of the life cycle of *Trypanosoma brucei.* This schematic traces the parasite from its infection of a human following a blood meal by a tsetse fly to its differentiation into bloodstream trypomastigotes and later its lifecycle inside the fly where it matures into procylic trypomastigotes, then epimastigotes and eventually the human-infective metacyclic trypomastigotes. Image courtesy of the United States Centers for Disease Control and Prevention (http://www.dpc.cdc.gov/dpdx).

prolonging host survival and increasing the probability of transmission of synchronous, developmentally co-ordinated parasites into the fly (Matthews, 2005; Matthews *et al.,* 2004). The mitochondrial activities of stumpy forms are also elaborated in order to preadapt to the glucose-depleted environment of the fly midgut (Matthews *et al.,* 2004).

Once taken into the fly, these stumpy forms differentiate into proliferative, procyclic parasites in the fly midgut with the kinetoplast positioned sub-terminally or midway between the nucleus and the posterior pole of the cell (Fig. 1.2). Slender forms are either lysed by lectins or differentiate into procyclic cells via the stumpy stage (Sbicego *et al.,* 1999). Procyclic cells have a different coat of GPI-anchored glycoproteins composed of EP and GPEET procyclins (named for the single letter code of a repeated amino acid motif). Procyclic cells have a fully active mitochondrion where substrate level phosphorylation is responsible for harnessing energy in the presence of glucose and relies on amino acid catabolism in its absence with a preference for *L*-proline (Marr *et al.,* 1977; Cazzulo *et al.,* 1985; Lamour *et al.,* 2005). Additionally, the cell length increases, particularly the distance between the kinetoplast and the posterior end of the cell body, during differentiation from bloodstream to procyclic form via active microtubule extension at the posterior end of the parasite (Sherwin *et al.,* 1987; Matthews *et al.,* 1995).

Procyclic cells migrate to the salivary glands from the fly midgut via the proventriculus, where they differentiate into the epimastigote forms- the sexual stage- via an asymmetric cell division resulting in parasites of two lengths (Lewis & Langridge,

1947; Van den Abbeele *et al.,* 1999; Sharma *et al.,* 2008a). While the long epimastigotes are terminal, the short epimastigotes are able to proliferate and are attached to the salivary gland epithelium via junctional complexes on the flagellar membrane (Tetley & Vickerman, 1985; Sharma *et al.,* 2008a). At this stage, the kinetoplast is anterior to the central nucleus. Epimastigotes express a stage-specific GPI-anchored surface protein family called *brucei* alanine rich proteins (BARP; Urwyler *et al.,* 2007).

Eventually, the parasites differentiate into a non-proliferative form called metacyclic forms, which reacquire a VSG coat and repress their mitochondria in preparation for the mammalian bloodstream milieu (Turner *et al.,* 1988). In the tsetse fly, the parasite development takes 20-30 days to reach maturation and generate the infective metacyclic forms (Aksoy *et al.,* 2003).

Of these morphologically distinct forms, only two can be cultured *in vitro* and subjected to genetic manipulation in the laboratory: the long-slender bloodstream and procyclic (fly midgut) trypomastigote forms (and to some extent, the stumpy forms).

1.1.4 Cell cycle of procyclic *T. brucei*

The single-copy organelles in the trypanosome cell (such as the kinetoplast, flagellum, mitochondrion, nucleus) must be faithfully duplicated and segregated prior to cell division. Figure 1.2 summarizes the temporal order of the key events in the procyclic cell division cycle with regard to the kinetoplast (mitochondrial DNA network) and nucleus. Both of these DNA-containing "organelles" have periodic DNA-synthesis

phases [nucleus (S_N) or kinetoplast (S_K)] and genome segregation events [mitosis (M) or kinetoplast segregation ($Segregation_K$)], and therefore broadly follow the scheme characteristic of the typical eukaryotic cell cycle –G_1, S, G_2, and M phases (Woodward & Gull, 1990).

Cells early in the cell cycle (G1/S) have a single nucleus and kinetoplast, and are, therefore, termed 1K1N cells. The first detectable cytological event in the cell cycle is the maturation of the pro-basal body into a basal body almost coincidentally with the initiation of S_N phase. The newly matured basal body elongates to subtend to a daughter flagellum that is attached to the old flagellum at the flagella connector (Sherwin & Gull, 1989a). The S_K phase is shorter than the S_N phase, possibly due to its smaller genomic contents and commences just prior to the S_N phase. While the flagellum grows, the kinetoplast segregates and moves apart within the cell. The kinetoplast is segregated and divided prior to the onset of mitosis, which produces late G2 cells with two kinetoplasts and one nucleus (2K1N). Mitosis is then initiated and at late anaphase/telophase, 2K2N cells are produced, positioning each daughter nucleus in the gap between the segregated basal bodies/kinetoplasts (Robinson *et al.,* 1995). Cytokinetic abscission occurs unidirectionally (anterior to posterior) along the helical axis of the cell to produce 1K1N daughter cells, containing single copies of the nucleus, kinetoplast and flagellum (reviewed in Farr & Gull, 2012).

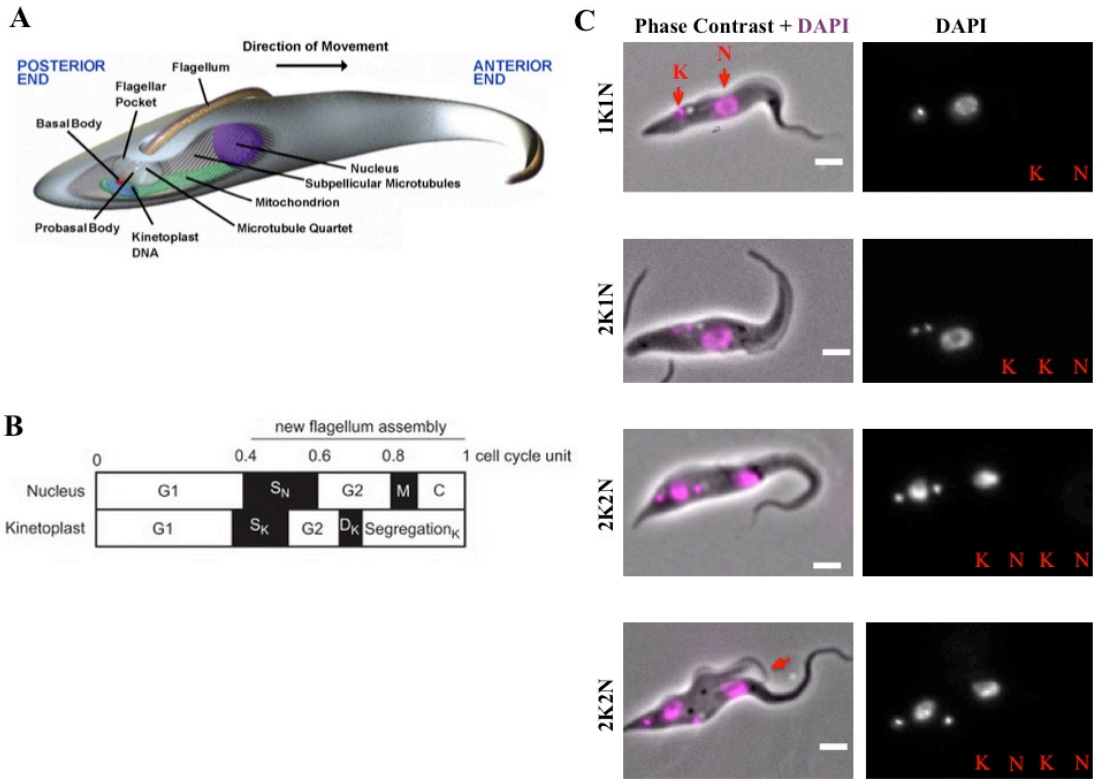


Figure 1.2. Cell cycle of procyclic *T. brucei.* A. A cartoon schematic of a procyclic *T. brucei* cell with key organelles and cytoskeletal structures labelled, adapted from Vaughan & Gull, 2008. **B.** A table summarizing the key events in the cell division cycle of a procyclic cell in terms of kinetoplast and nucleus. DNA-synthesis phases [nucleus (S_N) or kinetoplast (S_K)];M, mitosis; $Segregation_K$, kinetoplast segregation; C, cytokinesis. 1 cell cycle unit ≈ 8.5 hours for exponentially growing procyclic trypanosomes. Figure adapted from Olego-Fernandez *et al.,* 2009 and McKean, 2003. **C.** Phase contrast and DAPI micrographs of procyclic cells as they progress through the cell cycle as predicted by kinetoplast (K) and nucleus (N) configurations. Red arrow marks the position of the cleavage furrow. Bar, 3 μm.

With the temporal order of the key events in the procyclic cell division cycle characterized, one can readily determine the approximate position in the cell cycles of individual cells from asynchronous procyclic cultures by DAPI staining of the kinetoplast and nuclear DNA (Fig. 1.2C; Woodward & Gull, 1990).

The cell cycle of the bloodstream forms have not been studied at the same levels of detail as that of the procyclic cells. While they do share many similarities, profound differences exist in the regulation of their cell cycles and cytokineses (reviewed in McKean, 2003; Hammarton, 2006; Li, 2012). For example, depletion of a mitotic cyclin called CYC6 resulted in a mitotic block in both procyclic and bloodstream cells (Hammarton *et al.,* 2003). However, RNAi against *CYC6* generated anucleate cells with a single kinetoplast (called 'zoids', 1K0N) in procyclic cells and multikinetoplast cells with one nucleus in bloodstream cells. This suggested major differences in the cell cycle controls between the two developmental forms. While procyclic cells were able to undergo cytokinesis regardless of whether mitosis was complete, mitotic block in bloodstream form cells inhibited cytokinesis but not kinetoplast replication/segregation or an additional round of nuclear DNA synthesis (Hammarton *et al.,* 2003).

1.1.5 Cytoskeleton

During their life and cell cycles, trypanosomes undergo extreme changes in their cellular morphologies; all of which must occur within the borders of the microtubule-based cytoskeleton of the parasite. Essential microtubule-based cytoskeletal structures

include the flagellar axoneme, basal bodies, intranuclear mitotic spindle, and the sub-pellicular microtubule corset, which is a highly ordered cage lying just beneath the plasma membrane (Fig. 1.2A; reviewed in Gull, 1999). This mono-layered microtubule array consists of an array of >100 parallel and equally spaced microtubules -24 nm in diameter- that are cross-linked extensively, with each other and also with the overlying cell membrane (Sherwin & Gull, 1989a; Vickerman & Preston, 1976; Hemphill *et al.,* 1991).

Microtubules have an intrinsic polarity as they are made up of protofilaments containing α/ß tubulin heterodimers (Kohl & Gull, 1998). As such, all the microtubules of the *T. brucei* sub-pellicular cytoskeleton are arranged such that the plus or the dynamic ends are at the posterior end of the cell (Robinson *et al.,* 1995). However, the microtubules in the flagellar axoneme are arranged with the plus ends at the anterior end of the cell. In trypanosomes, the sub-pellicular microtubule array remains intact during cell division. During the early stages of the cell cycle, the posterior ends elongate due to both microtubule extension and the intercalation of new microtubules in the microtubule array such that each daughter cell inherits a complete microtubule corset upon cytokinesis in a semi-conservative fashion (Sherwin *et al.,* 1987; Gull, 1999).

1.1.5.1 Microtubule-associated proteins

The cross-linkages in the sub-pellicular cytoskeleton are mediated by a number of microtubule-associated proteins (MAPs), which regulate inter-microtubule interactions

and/or membrane-cytoskeleton linkages. Such cross-linking must contribute to the high stability of this array as treatment by a non-ionic detergent produces cytoskeletons in contrast to the microtubular structures in mammalian cells, which depolymerize readily upon cell disruption (Sherwin & Gull, 1989a; Hemphill *et al.,* 1991). Additionally, the sub-pellicular corset fractionates with the plasma membrane suggesting their stable association (Gull, 1999).

Several MAPs have been identified yet the molecular mechanisms of their functions or their targeting remain poorly understood. Examples of MAPs include cytoskeleton-associated protein 5.5 (CAP5.5), its bloodstream-specific variant, CAP5.5V, the 41kDa protein p41, the differentially expressed CAP15, and CAP17 (Olego-Fernandez *et al.,* 2009; Hertz-Fowler *et al.,* 2001; Schneider *et al.,* 1988; Vedrenne *et al.,* 2002).

CAP5.5 was identified as a result of antibodies raised against the whole cytoskeleton, and its in-paralogous variant CAP5.5V identified following the genome sequencing of *T. brucei* (Birkett *et al.,* 1992; Matthews & Gull, 1994; Berriman *et al.,* 2005). Both have been shown to associate with the sub-pellicular cytoskeleton by biochemical and structural analyses although they are not found on the axonemal microtubules of the flagellum (Olego-Fernandez *et al.,* 2009; Hertz-Fowler *et al.,* 2001). They are hypothesized to mediate cross-linking between the cytoskeleton and the overlying plasma membrane, as CAP5.5 has been biochemically demonstrated to be dually acylated with myristic and palmitic acids (Hertz-Fowler *et al.,* 2001). Protein

lipidation, especially dual acylation, is exclusively found in membrane-associated proteins and are thought to mediate high-affinity interactions with membranes (Resh, 1999). CAP5.5 and CAP5.5V are essential and required for the correct cell morphogenesis of the procyclic and bloodstream forms respectively, with abnormalities in the sub-pellicular microtubule corset organization observed, among other phenotypes, following RNAi against either calpain. CAP5.5/V will be discussed in greater detail in Chapter 1.2.7.

Like CAP5.5 and CAP5.5V, p41 is thought to mediate membrane-microtubule interactions as it has been shown to be dually acylated *in vivo* and binds to isolated cytoskeleton and purified microtubules *in vitro* in a calcium-dependent manner in procyclic cells (Schneider *et al.,* 1988). CAP15/17, on the other hand, are differentially expressed like CAP5.5/V, with CAP15 expression more abundant in bloodstream forms and CAP17 expressed only in procyclic trypanosomes (Vedrenne *et al.,* 2002). Immunofluorescence studies have shown both proteins to associate only with the anterior part of the sub-pellicular microtubule corset and to stabilize microtubules when expressed in mammalian cells. Upon over-expression of either protein in procyclic parasites, the proteins were redistributed across the cell body and caused organelle segregation defects, possibly by stabilizing the dynamic posterior ends of the cytoskeleton (Vedrenne *et al.,* 2002). However, not all MAPs are dually acylated like CAP5.5/V and p41 or differentially expressed like CAP5.5/V and CAP15/17. For examples of other MAPs, please refer to these reviews (Kohl & Gull, 1998; Portman & Gull, 2012).

1.1.6 Gene Expression

In trypansomatid parasites such as *T. brucei*, *T. cruzi* and *Leishmania* spp., the genomes are organized in tandem, polygenic clusters of functionally unrelated genes that are transcribed as polycistronic units and then, processed into individual mRNAs (Berriman *et al.,* 2005). Functional and stable mRNAs are created by 5' and 3' modifications similar to other eukaryotic mRNAs. However, in trypanosomatids and a few other eukaryotic species, while the mature mRNAs are polyadenylated at the 3' end, they are capped at the 5' end by spliced leader (SL) *trans*-splicing where the 5' terminal (~40 nt) of the SL RNA sequence is fused onto the 5' end of each mRNA (Ullu *et al.,* 1993; Günzl, 2010). While these polygenic clusters are largely transcribed by RNA polymerase II, RNA polymerase I transcribes rRNA and mRNA encoding surface proteins of the parasite such as the VSG and EP/GPEET (reviewed in Palenchar & Bellofatto, 2006).

Given the polycistronic organization of the trypanosomatid genome, most of the regulation is largely post-transcriptional where mRNA levels are determined by modulating mRNA stability and processing. Such regulation is mediated by *trans*-acting factors that recognize *cis*-acting sequences such as the 3' UTR of mRNAs (Clayton, 2002; Ouellette & Paladopoulou, 2009). Another method to regulate gene expression is epigenetic. For instance, transcription of reporter genes under the control of either *rRNA* or *vsg* expression site promoters were shown to be differentially repressed depending on

their proximity to telomeres or silenced genes (Horn & Cross, 1997). For more detail, please refer to these reviews (Figueiredo *et al.,* 2009; Alsford *et al.,* 2012).

1.2 Calpains

Calpains (EC 3.4.22.17; Clan CA, family C02) are a group of multi-domain intracellular calcium-dependent cysteine proteases found in most eukaryotes. Genome databases revealed no archaeal calpains but a few bacterial sequences related to the catalytic core domain II of calpains have been found (Croall & Ersfeld, 2007). These are not promiscuous degradative proteases but instead cleave a restricted set of protein substrates upon calcium activation (Croall & Ersfeld, 2007). Substrates such as talins and spectrins implicate calpains in cytoskeletal remodelling in addition to a range of other physiological processes including signal transduction, muscle function, cell motility, apoptosis, cell cycle regulation, differentiation and development (Lebart *et al.,* 2006; Goll *et al.,* 2003). In humans, elevation of calpain activity or prolonged calpain activation can cause off-target proteolysis and as such, has been correlated with Alzheimer's disease, type II diabetes, myopathy, cardiac ischemia and rheumatoid arthritis (reviewed in Goll *et al.,* 2003). In many animal disease models, administration of calpain inhibitors has been demonstrated to ameliorate the aforementioned conditions. Therefore, their physiological roles, and in particular, their contribution to human diseases have been extensively studied. No calpains have been as exhaustively characterized as the two isoforms, calpain 1 and 2, which are the most abundant calpains

found in almost all tissues and cell types of mammals. As such, the focus of the following analyses of calpain structure and function shall be on these two prototypical calpains.

1.2.1 Typical calpains – structure

Calpains can be broadly categorized into typical and atypical calpains, with the distinction being the presence of a calcium-binding C-terminal domain in typical calpains. While, typical calpains are largely restricted to animals (metazoans), atypical ones can be found in plants, animals, fungi, and bacteria (Sorimachi *et al.,* 2011a; Sorimachi *et al.,* 2011b). Calpains 1 (CAPN1) and 2 (CAPN2) are typical calpains. In humans, nine out of the 15 independent calpain genes encode typical calpains (CAPN 3, 8, 9, 11, 12, 13, 14) with one (CAPN4) encoding the small subunit (Goll *et al.,* 2003). Some of these are ubiquitously expressed in all tissues (CAPN1, 2, 4, 5, 7, 10, 12, 14, 15) while the rest are tissue-specific. Typical calpains are dimers consisting of a large 80 kDa catalytic subunit with four domains (I-IV) and a smaller 29 kDa regulatory subunit with two domains (V-VI). The C-terminal domain IV mediates heterodimerization of both CAPN1 and CAPN2 with their smaller subunits or homodimerization as in the case of CAPN3 and CAPN13 (Ravulapalli *et al*., 2009).

In 2012, calpain biologists proposed a new and unified nomenclature system for calpain domains to reflect the growing body of structural characterization of each domain (Campbell & Davies, 2012; Ono & Sorimachi, 2012). This new system of domain architecture and boundaries of calpains is shown with an example of the old system, as

well as a calcium-free crystal structure of human CAPN2 (Hosfield *et al.,* 1999; Strobl *et al.,* 2000; Fig. 1. 3).

The new classification system still divides typical calpains into six domains/regions. **(1). The N-terminal anchor helix** (23 aa) forms a single α-helix and contributes to protein stability by forming associations with domain VI of the small subunit [now called PEF (S)]. Originally, this α-helix was grouped as part of domain I (1-80). However, x-ray crystallography revealed that domain I was not a structural entity in its own right. For while the N-terminal α-helix (1-23) was anchored into the calpain small subunit, the rest packed against the adjacent protease core (Hosfield *et al.,* 1999). Therefore, those C-terminal amino acids were grouped into the protease core (Fig. 1.3).

(2). The protease core domain (PC1 and PC2) contains two subdomains, PC1 (186 aa) and PC2 (131 aa). These subdomains were originally referred to as domain II (Fig. 1.3A). Homology to this domain is often used as criterion for designation as a calpain (for exceptions see section on kinetoplastid calpains). The protease core domain contains the catalytic triad residues characteristic of cysteine proteases: cysteine (C115 and C105 in CAPN1 and CAPN2 respectively), histidine (H272, H262) and asparagine (N296 and 286). While, both H-N residues are on the PC2 subdomain, C is found on the PC1. Unlike most allosterically regulated enzymes where enzyme activation involves the relieving of steric hindrance to a pre-formed active site, calpain activation requires the conformational alignment of the two PC1 and PC2 subdomains to generate the active site

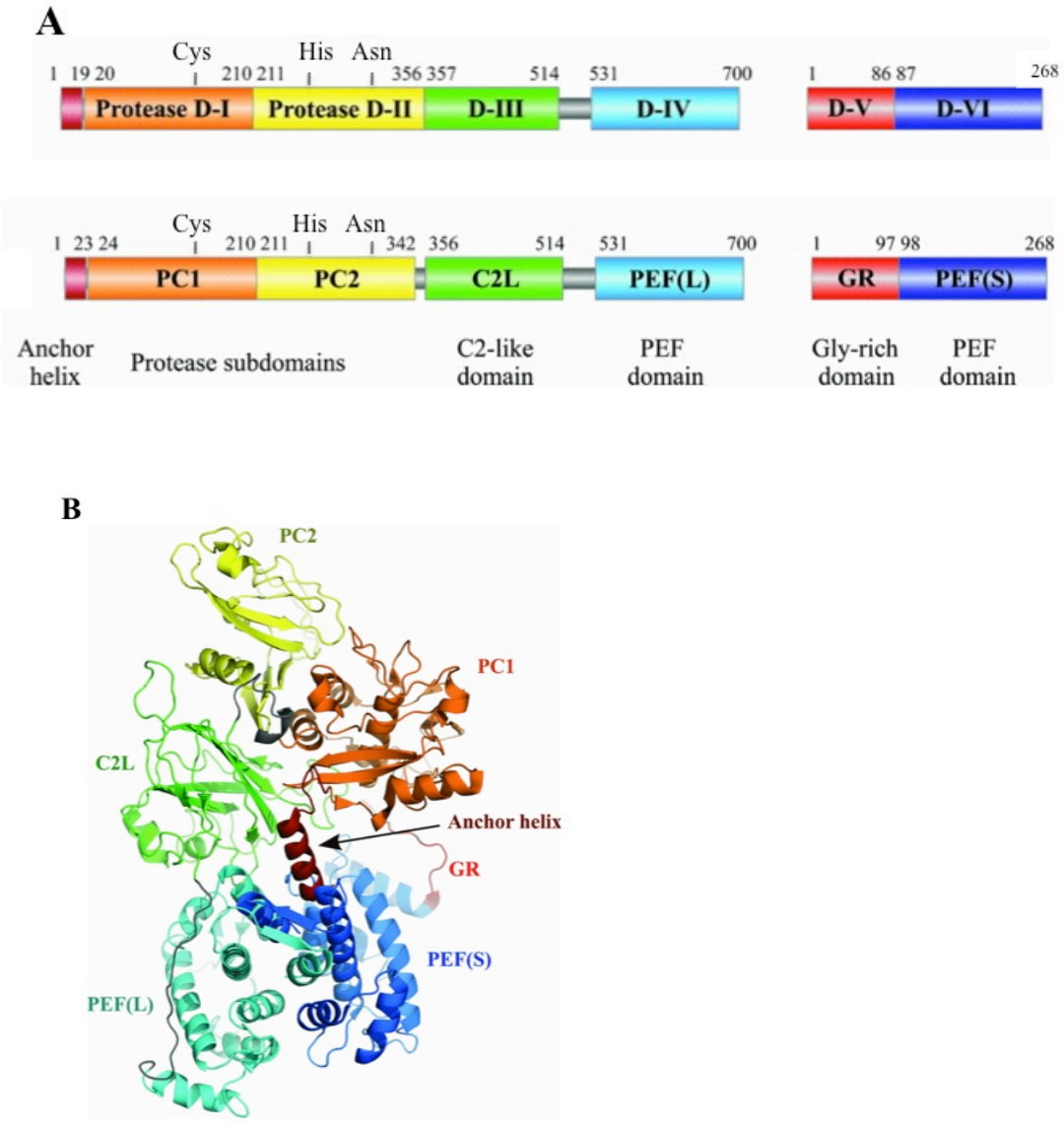


Figure 1.3. Domain architecture, nomenclature and structure of typical calpain, CAPN2. A. The domain architecture of typical calpains as exemplified by rat CAPN2 is shown along with the old (top) and new (bottom) nomenclature system for the various domains. **B.** Crystal structure of human CAPN2 is shown in a calcium free (inactivated) state (PDB code 1KFU). The human CAPN2 is shown instead of rat CAPN2 owing to the omission of several regions during the latter's structural characterization. The grey space indicates linker regions. Images courtesy of Campbell & Davies, 2012.

(Croall & Ersfeld, 2007; more on calpain activation in 1.2.2). When crystallized, the protease core domain yielded a calcium-bound structure with one calcium atom bound to each PC subunits (Moldoveanu *et al.,* 2002). Further characterization revealed the participation of five acidic residues in binding the divalent ions (Moldoveanu *et al.,* 2004a).

Interestingly, the protease core domain of rat CAPN1 was shown to have weak but detectable proteolytic activity upon calcium loading. This was observed even in the absence of the other domains, suggesting that the protease core acts co-operatively upon calcium activation to form an independent and functional protease unit (Moldoveanu *et al.,* 2002; Moldoveanu *et al.,* 2004a). Comparison of the proteolytic activity of the protease core and the whole calpain revealed that while both enzymes had similar initial reaction rates, the protease core maintained its enzymatic activity for a longer duration (Campbell & Davies, 2012).

(3). C2-L domain (158 aa; previously known as domain III) is a C2-like domain with an acidic loop constituted by a stretch of consecutive aspartic acid and glutamic acid residues (Fig. 1.3). C2 domains are found in various proteins such as protein kinase C and phospholipase C and are known to bind to Ca^{2+} and phospholipids (Rizo, J. *et al.,* 1998). It has been suggested that the acidic loop facilitates a calcium-dependent interaction with the phospholipid membrane and/or regulation of calpain activity (Fernandez-Montalyan *et al.,* 2004).

(4). (5). The C-terminal domains of the large and small subunits are called **PEF (L)** and **PEF (S)** with PEF referring to the penta EF-hand motifs that are found on both subunits (Fig. 1.3). The parenthetical letter refers to whether that domain belongs to the large (L) or small (S) subunit of the dimeric calpain. They were called domain IV and VI respectively before. The two domains, approximately 170 amino acids each, share an approximate 45% sequence identity and are homologous to calmodulin (Strobl *et al.,* 2000). Four out of the five EF-hands are known to bind calcium in both domains, with the remaining EF hand mediating dimerization between the big and small subunits (Kretsinger, 1997).

(6). GR domain is the N-terminal domain (95 aa) of the small subunit, originally called domain V (Fig 1.4). This domain contains a stretch of glycine residues at the N-terminus (40 glycines total) and most likely penetrates the membrane (Dennison *et al.,* 2005). This domain is highly flexible and structurally unresolved.

1.2.2 Typical calpains – function

Typical calpains are not completely degradative proteases but instead cleave a restricted set of protein substrates upon calcium activation to effect a change in the substrate's function (Croall & Ersfeld, 2007). Their activity is tightly regulated *in vivo* by its endogenous inhibitor, calpastatin and the cytosolic concentration of free calcium (Suzuki *et al.,* 1987). More than 100 calpain substrates have been identified which were largely transcription factors, signalling enzymes and cytoskeletal proteins (Franco *et al.,*

2005). While these substrates have helped one understand the roles of calpains in cellular processes and diseases, their specific mechanisms, *in vivo* regulation, and physiological substrates remain largely unknown (Goll *et al.,* 2003).

One of the better-characterized physiological roles of calpains is their involvement in cytoskeletal remodelling processes such as cell motility in animal cells (Franco *et al.,* 2005; Lebart *et al.,* 2006; Goll *et al.,* 2003). Cell migration involves a series of well-coordinated steps: (1). In response to certain stimuli or cues, the 'front' of a cell forms a membrane ruffle following actin polymerization. (2). That protrusion is stabilized by an interaction with the extracellular matrix (ECM) through focal adhesion complexes, linking the ECM to the actin cytoskeleton. (3). There is an active reorganization of the actin cytoskeleton to move forward. (4). The cytoskeletal anchorages and adhesion complexes on the rear end must be disassembled to allow directional movement (Chan *et al.,* 2010; Franco *et al.,* 2005; Lebart *et al.,* 2006). It has been shown that limited proteolysis of cytoskeletal proteins such as talin1, focal adhesion kinase (FAK), other adhesion complex proteins and migration-related proteins plays a critical role in adhesion disassembly during cell migration (Chan *et al.,* 2010). For instance, RNAi-based knockdown of CAPN2 in human fibroblasts resulted in decreased proteolysis of cytoskeletal proteins like FAK, talin 1, spectrin etc (Franco *et al.,* 2005). Additionally, embryonic fibroblasts from CAPNS1 knockout mice (CAPNS1 referring to the small subunit of CAPN1) also showed similar results (Dourdin *et al.,* 2001).

Therefore, it seems that calpains, such as CAPN2, regulate the proteolysis of several substrates that contribute to cell motility and migration.

While we have some idea of the pleiotropic roles of calpains in human physiology, a more detailed understanding of how they actually work may be beneficial. One of the better models for calpain activation came from comparative analyses of the structures of CAPN2 in active calcium-bound or inactive calcium-free states (Hosfield *et al.,* 1999; Strobl *et al.,* 2000). While calcium loading does not change the secondary structures of the various domains as determined by biophysical analyses, binding of the two calcium ions to PC1 and PC2 sub-domains creates a conformational rearrangement that reduces the distance between the catalytic triad residues from 8.5 Å to around 3.5 Å, adequate for nucleophilic charging of the catalytic cysteine. Additionally, this reorientation enables the PC domain to create an active site cleft. The other eight calcium ions that bind to the PEF domains of both the large and small subunits are thought to provide an added level of regulation to calpain activity and do not necessarily act as primary regulators of calpain activity.

1.2.3 Atypical calpains

While typical calpains are largely restricted to animals (metazoans), atypical ones can be found in plants, animals, fungi, and bacteria (Sorimachi *et al.,* 2011a; Sorimachi *et al.,* 2011b). In fact, atypical calpains are far more widespread and numerous than typical calpains, which has led some to question the correctness of the terms “atypical” for these

proteins (Campbell *et al.,* 2012). Among the fifteen mammalian calpains, five of them are classified as atypical: 5, 6, 7, 10 and 15 (Goll *et al.,* 2003). Despite being far more numerous than typical calpains, no unified classification of atypical calpains exists as a variety of domain and motifs combinations have been found including an additional C2-like domain, zinc fingers, microtubule interacting and transport (MIT) domains (Joyce *et al.,* 2012).

Atypical calpains are those that lack the PEF (L) domain (also known as domain IV) and as such, are predicted to be monomeric (Pietsch *et al.,* 2010). The loss of a domain characteristic of typical calpains does not necessarily equate being an inactive cysteine protease as atypical calpains, such as the maize calpain defective kernel 1, have been demonstrated to be calcium-regulated proteases (Wang *et al.,* 2003). All atypical calpains share homology to the catalytic core domain of typical calpains (>20-25% amino acid identity), the basic for classification into the calpain superfamily (Wilson *et al.,* 2000). These atypical calpains may deviate from the canonical Ca^{2+} dependence for activation such as CAPN7 or may lack the canonical catalytic triad residues like CAPN6, but atypical calpains have evolved to assume a range of roles that are as diverse as they are ubiquitous (reviewed in Joyce *et al.,* 2012). As such, atypical calpains have been implicated in a range of physiological processes such as myofilament turnover, neurodegeneration, sex determination, and adaptive responses to alkaline environments although the molecular mechanisms are largely unknown (Ono & Sorimachi, 2012).

One of the better-characterized atypical calpains in mammals is the ubiquitously expressed CAPN7 - the study of which may offer one perspectives into the evolution of atypical calpains in general. CAPN7 possesses a tandem repeat of microtubule-interacting and trafficking molecule domains (MIT) at its N-terminus and a C-terminal domain that shares sequence homology to the C2-L domain (domain III) of typical calpains. It has been proposed that a tandem duplication of a C2-L domain-containing DNA in an ancestral calpain gene gave rise to this domain, suggesting perhaps a functional conservation between it and the C2-L domain (Franz *et al.,* 1999). CAPN7 has no calcium-binding sites on all its domains but does possess the canonical catalytic triad residues, suggesting a calcium-independent regulation.

Initial studies revealed no *in vitro* proteolytic activity for recombinant CAPN7 in COS cells against universal protease substrates such as casein or against synthetic peptide substrates (Futai *et al.,* 2001). However, the Maki lab recently demonstrated CAPN7's autolytic activity but showed that it is dependent on MIT interactions with MIT-interacting motifs of other proteins in a calcium independent manner (Yorikawa *et al.,* 2008). This suggests that CAPN7 might have evolved from calcium dependence of ancestral (typical) calpains to later mediate its activity by interactions of its MIT domains with other MIT-interacting complexes.

MIT domain-containing proteins have been known to interact with proteins involved in endosomal sorting complex required for transport (ESCRT)-III (Ghazi-Tabatabai *et al.*, 2009). One of the members of the ESCRT-III family is IST1, which has

MIM -MIT-interacting motifs- that facilitate interactions with MIT domains such as those of CAPN7 (Osako *et al.,* 2010). This suggests that CAPN7 may be an active protease with a role in the ESCRT system. However, fungal and yeast CAPN7 orthologs do not have MIT domains, implying CAPN7's evolution from lower eukaryotic calpains to execute a physiological role that utilizes ESCRT interactions for activation of its activity (Franz *et al.,* 1999; Denison *et al.*, 1995).

1.2.4 Kinetoplastid calpains

Systematic analyses of the genomes of three kinetoplastid parasites, *Trypanosoma brucei, Trypanosoma cruzi* and *Leishmania major* revealed the presence of a large number of calpain-like proteins, all of which lacked the PEF domain characteristic of typical calpains (Ersfeld *et al.,* 2005). Such *in silico* analyses were accomplished by "baiting" the domain II of a previously characterized calpain-like protein, CAP5.5, into genome databases to uncover 12 calpain-related sequences in *T. brucei,* 15 in *T. cruzi* and 17 in *L. major.* The domain II of these calpains shared a 25% overall sequence identity compared with the corresponding catalytic domains (PC) of typical calpains and a 45% similarity based on the physicochemical properties of amino acids. A majority of the kinetoplastid calpains containing domain II also showed some degree of similarity (15-23% identity; 25-35% similarity) between their domain IIIs and the C2-L domain (III) of typical calpains.

Upon further analyses of the retrieved calpain sequences, it was observed that the domain (~100 aa; termed domain I^K) N-terminal to domain II shared significant similarities to each other and was found either as "stand-alone" sequences or fused to domain II-containing calpains. Domain I^K shared no similarities to any other sequences in the database other than to the kinetoplastid calpain-like sequences. Therefore, these "stand-alone" sequences were included as calpain-like genes, increasing the number of calpains to 18 in *T. brucei*, 24 in *T. cruzi* and *27 in L. major*. This is in stark contrast to the number of calpain-like genes (1 each) found in other protozoan parasites such as *Plasmodium falciparum*, *Cryptosporidium parvum, Theileria annulata* and *Entamoeba histolytica* (Gardner *et al.,* 2005; Anderson & Loftus, 2005; Abrahamsen *et al.,* 2004; Croall & Ersfeld, 2007). To my knowledge, this represents the largest number of calpains (typical or atypical) in any sequenced organism containing calpain-encoding genes.

In the three kinetoplastid parasites, a majority of these calpains contain mutation(s) in their catalytic triad residues except for 1 (out of 18) in *T. brucei,* 4 (out of 27) in *L. major* and 2 (out of 24) in *T. cruzi*. However, there is only one kinetoplastid calpain that contains not only the catalytic triad residues characteristic of typical calpains but also the conserved calcium-binding sites found in calcium-regulated cysteine proteases (LmjF25.1480). However, whether any of these calpains, even those with mutations in their catalytic triads, have intrinsic proteolytic functions remains unknown (for more, see chapter 1.2.7). Additionally, 24 out of the 69 identified calpains contain a kinetoplastid calpain-specific fatty acid acylation motif, which are predicted to mediate

membrane associations (Resh, 2006; Ersfeld *et al.,* 2005; more on acylation in chapter 1.2.6).

These kinetoplastid calpains have been classified on the basis of their predicted domain architecture into five groups (Fig. 1.4; Ersfeld *et al.,* 2005). The classification system used for kinetoplastid calpains differs from that adopted for typical calpains owing to the presence of kinetoplastid-specific domains, and lack of functional or structural information regarding these domains. In all the groups, the C-terminal domain was named C domain in order to highlight its difference from the penta-EF hand domain of typical calpains. This C domain has no homology to any sequence in the databases and does not contain any recognizable motifs. Additionally, that N-terminal domain that was highly conserved in kinetoplastids was named domain I^K in order to distinguish this domain from a heterogenous domain (I^H) with little homology to domain I^K. Domain I^K contains the acylation motif.

Group **1** (I^K-II-III-C) calpains contain sequences with a four-domain architecture resembling that of the large subunit of typical calpains, with the obvious difference being their N- and C-terminal domains (named I^K and C respectively). Proteins in this group are hypothesized to be acylated as has been experimentally demonstrated for group 1 calpains, such as *T. brucei* CAP5.5 (Hertz-Fowler *et al.,* 2001).

Group **2** (I^H -II-III-C) is similar to Group **1**, except for its N-terminal domain as group 2 calpains contain the non-conserved heterogeneous domain I^H (Fig. 1.4).

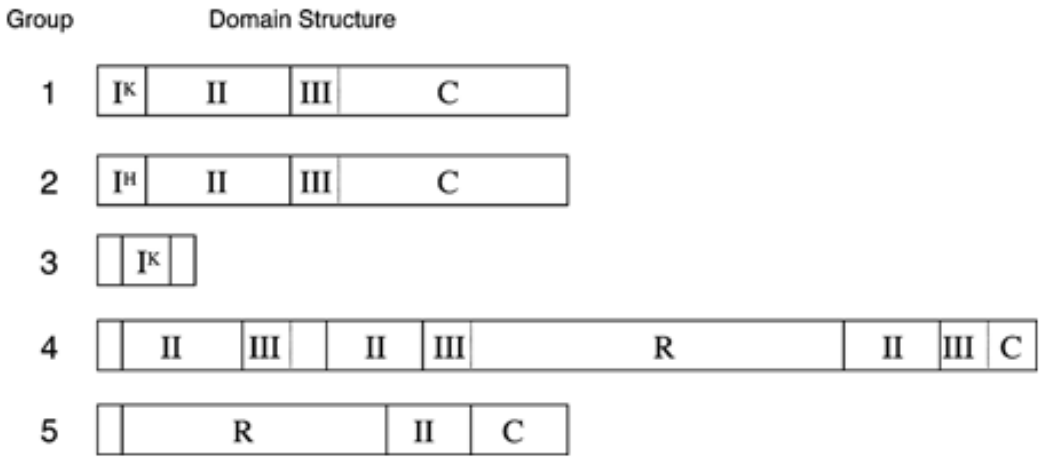


Figure 1.4. Classification and domain architecture of kinetoplastids calpains. I^K, kinetoplastid-specific N-terminal domain; I^H, heterogeneous N-terminal domain; C, C-terminal domain; R, repeats. Classification and image courtesy of Ersfeld *et al.,* 2005.

Group **3** (I^K) sequences contain a single domain that is highly similar to the N-terminal domain of group 1 sequences (Fig. 1.4). In *L. major*, proteins in this group are called small myristoylated proteins (SMPs) whereas in *T. brucei* and *T. cruzi*, they are called small kinetoplastid calpain-related proteins (SKCRPs).

Group **4** (II-III-II-II-repeats-II-III-C) proteins contain three repeats of domains II and III, with tandem repeat sequences of ~70 amino acids preceding the third repeat of domain II and III. It has been hypothesized that these internal repeats mediate interactions with microtubules as they have been observed in a number of cytoskeleton-associated proteins in trypanosomatids (Ersfeld *et al.,* 2005; Gull, 1999).

Group **5** (repeats-II-C) calpains contain tandem repeats like those seen in group 4 followed by domain II and the C-terminal domain.

1.2.5 Localization and biochemical characteristics of kinetoplastid calpains

The cellular localization and biochemical characteristics particularly the status of fatty acid acylation of several kinetoplastid calpains have been characterized (summarized in Table 1.1). While one may notice that the majority of these calpains are acylated (mono or otherwise) and are associated to cytoskeletal structures and the membrane, these are not necessarily prototypical characteristics of kinetoplastid calpains. For those *T. brucei* calpains denoted as localizing to the cell body, the authors did not determine whether they associated with the sub-pellicular cytoskeleton (Liu *et al.,* 2010).

Table 1.1. Biochemical features and localization of the hitherto characterized kinetoplastid calpains.

Name	GeneDB ID	Catalytic triad (Group #)	Acylation motif	Localization (M=membrane association)	Ref.
Trypanosoma brucei					
TbCALP1.3	Tb927.1.2120	YHN (1)	PA*	Flagellar tip (M)	[1]
CAP5.5	Tb927.4.3950	SYN (1)	PA*; MA*	Subpell. cyto. (M)	[2]
TbCALP4.2	Tb927.4.3940	SSN (1)	None	Cell body (M)	[1]
CAP5.5V	Tb927.8.8330	SYN (1)	PA; MA	Subpell. cyto. (M)	[1, 3]
TbCALP6.1	Tb927.6.3310	SHS (2)	None	Cell body (M)	[1]
SKCRP1.4	Tb927.1.2150	n/a (3)	MA	Cell body (M)	[1]
SKCRP1.5	Tb927.1.2160	n/a (3)	PA; MA	Flagellum (M)	[1]
SKCRP1.6	Tb927.1.2230	n/a (3)	PA*; MA	Cell body (M)	[1, 10]
SKCRP7.1	Tb927.7.4060	n/a (3)	None	Cell body	[1]
SKCRP7.2	Tb927.7.4070	n/a (3)	None	Cell body & flagellar tip	[1]
GM6	Tb11.47.0036	YCC (5)	None	FAZ	[4]
Trypanosoma cruzi					
H49	Tc00.1047053506721.30	LCN/CQN/CCQN (4)	None	FAZ in epimastigotes	[5]
FRA	Tc00.1047053511441.10	Y-- (5)	None	Flagellar cytoskeleton	[6]
Leishmania major					
SMP-1	LmjF.20.1310	n/a (3)	PA*; MA*	Flagellum (M; promastigotes)	[7]
SMP-2	LmjF.20.1300	n/a (3)	MA*	Flagellar pocket (M; promastigotes)	[8]
SMP-4	LmjF.20.1280	n/a (3)	MA*	Pellicular membrane (promastigotes)	[9]

For acylation motif, only those calpains containing the dual acylation motif are indicated, with 'none' indicating the lack of a myristoylation motif and an asterisk (*) indicating biochemical confirmation of acylation. PA, palmitic acid; MA; myristic acid; CAP5.5/V, cytoskeleton-associated protein 5.5/V; SKCRP, small kinetoplastid calpain-related protein; SMP, small myristoylated protein; FRA, flagellar repetitive antigen; FAZ, flagellar attachment zone or complex connecting the cytoskeleton to the flagellum. [1], Liu *et al.,* 2010; [2], Hertz-Fowler *et al.,* 2001; [3], Olego-Fernandez *et al.,* 2009; [4], Muller *et al.,* 1992; [5], Galetović *et al.,* 2011; [6], Lafaille *et al.,* 1989; [7], Tull *et al.,* 2004; [8], Tull *et al.,* 2010; [9], Tull *et al.,* 2012. [10], Emmer *et al.,* 2011.

Additionally, several of the *T. brucei* calpains were differentially expressed in either procyclic or bloodstream forms as assessed by qRT-PCR. In addition to CAP5.5 and CAP5.5V, which have been shown previously to be procyclic and bloodstream-specific calpains respectively, TbSKCRP1.6 and TbSKCRP7.2 were found to be differentially expressed in procyclic cells (Liu *et al.,* 2010; Olego-Fernandez *et al.,* 2009).

1.2.6 Acylation

Dual acylation appears to be a kinetoplastid calpain-specific feature as only kinetoplastid calpains are modified with myristic acid (C14:0) and/or palmitic acid (C16:0), with the one known exception in the calpain world being the sole calpain in *P. falciparum*, *Pf*_calpain, which is however, only monoacylated with myristic acid (Russo *et al.,* 2009). As explained before, one of the major roles of metazoan calpains is cytoskeletal remodelling; interestingly, while these typical calpains are themselves not acylated, their cytoskeletal targets, such as spectrin, vinculin, ankyrin and band 4.1, are acylated and membrane-associated (Staufenbiel & Lazarides, 1986; Burn &Bruger, 1987; Maretzki *et al.,* 1990; Mariani *et al.* 1993; Bhatt *et al.,* 2002).

N-myristoylation is a co-translational process in which protein sequences with a N-terminal motif consisting of MGXXXS/T have myristic acid added to the glycine by an amide bond, following the removal of the N-terminal methionine (Towler *et al.,* 1987; Resh, 2006). This reaction is facilitated by a myristoly-CoA:protein N-myristoyltransferase that was found in both in the cytosolic and membrane fractions in *Leishmania* spp. and *T. brucei* (Price *et al.,* 2002). On the other hand, there is no

consensus sequence for palmitoylation but for dually acylated proteins containing the sequence Met-Gly-Cys at their N-terminus, Gly2 myristolyation facilitates the thioester bond linkage of palmitic acid to the neighbouring cysteine (Resh, 1999; Godsel & Engman, 1999; Mumby, 1994).

Kinetoplastid parasites (*T. cruzi*, *T. brucei*, and *L. major*) are predicted to express at least 60 myristoylated proteins, with small molecule inhibitors against the myristoylation enzyme being developed as future trypanocidal drugs in humans and cattle (Frearson *et al.,* 2010; Mills *et al.,* 2007). Myristoylation is known to regulate membrane targeting and trafficking, however, often is insufficient by itself unless accompanied by additional acylation events like palmitoylation and/or the presence of basic amino acids that electrostatically bind to the polar phospholipid heads of membranes (Resh, 2006). Fatty acid acylation is but one mechanism for trafficking proteins to the membrane.

In trypanosomatids, proteins with lipophilic moieties are predominantly targeted to the flagellar membrane, where there is an enrichment of sterols, glycolipids, sphingolipids, all of which are components of canonical lipid rafts in contrast to the microenvironment of the cell body (pellicular) membrane (Chailley, 1985; Tetley, 1986; Kaneshiro *et al.,* 1984). Association of dually acylated proteins with such lipid rafts have been suggested as a possible mechanism for flagellar membrane trafficking in kinetoplastids (Fridberg *et al.,* 2007; Tyler *et al.,* 2009). Numerous dually acylated proteins such as flagellar calcium binding protein (FCaBP) in *T. cruzi*, SMP-1 in *L. major*, and calflagins in *T. brucei* localize to lipid rafts/subdomains in their respective flagellar membranes, where the necessity and

importance of lipidation have been experimentally demonstrated for these proteins' localization (Maric *et al.,* 2011, Tull *et al.,* 2012; Tyler *et al.,* 2009). Cellular localization analyses of mono- and diacylated proteins suggest that while dual acylation in itself appears to elicit a flagellar membrane association in kinetoplastid parasites, this signal can be superseded by additional cues (Emmer *et al.,* 2009). For example, CAP5.5 and CAP5.5V associate with sub-pellicular cytoskeleton and are predicted to interact with the cytosolic face of the pellicular membrane but are excluded from the flagellum (Hertz-Fowler *et al.,* 2001; Olego-Fernandez *et al.,* 2009).

1.2.7 CAP5.5 and CAP5.5V

CAP5.5 was identified as a result of antibodies raised against the whole cytoskeleton, and its in-paralogous variant CAP5.5V identified following the genome sequencing of *T. brucei* (Birkett *et al.*, 1992; Matthews & Gull, 1994; Berriman *et al*., 2005). CAP5.5 was so named as immunolocalization studies showed it to localize to the sub-pellicular cytoskeleton, hence, cytoskeleton-associated protein 5.5 (5.5 for the clone from which it was discovered). CAP5.5V, which is found on the sub-telomeric region on chromosome 8, is thought to have occurred as a result of a recent chromosomal duplication event of 0.5 Mb of chromosome 4, which included the locus where CAP5.5 is found.

CAP5.5 and CAP5.5V are group 1 calpains, containing a four-domain architecture of I^K-II-III-C, with I^K being the kinetoplastid-specific N-terminal domain and C being the C-terminal domain. Domains II and III, like all kinetoplastid-

calpains, bear some sequence homology to the corresponding domains of typical calpains (for the sequence identity, see chapter 3.1). Both CAP5.5 and CAP5.5V are predicted to be dually acylated with myristic and palmitic acids, which are thought to mediate their associations with the pellicular membrane that might in turn confer these calpains some role in intracellular signalling and regulation like other acylated proteins. (Hertz-Fowler *et al.*, 2001; Resh, 2006). Membrane-cytoskeleton associations are hard to determine even by immunogold electron microscopy, owing to the distance (~10 nm) between the membrane and the cytoskeleton, and the inter-microtubule distance of 18-20 nm. However, owing to the ultrastructural, biochemical and fluorescence analyses of CAP5.5, which show it to associate with the sub-pellicular cytoskeleton and the dual acylation, CAP5.5's possible function may be to mediate the physical linkages between the membrane and the underlying cytoskeleton (Hertz-Fowler *et al.,* 2001).

The sequence identity between CAP5.5 and CAP5.5V for the first 720 amino acids of CAP5.5 is 84% and 75% across their full lengths, differing primarily in their C-terminal domain. The N-terminus of CAP5.5/V contains a proline-rich region immediately following the myristoylation-palmitoylation signal. The sequence PxxP conforms to the binding sequence motif found in many ligands of SH3-type domain containing proteins (Ren *et al.,* 1993; Yu *et al.,* 1994). Proline-rich regions are recognized as mediating protein-protein interactions, particularly at the plasma membrane interface (Buday *et al.,* 1999; Niebuhr *et al.,* 1997). This region may serve to stabilize the membrane associations. In domain II, while typical calpains contain the catalytic triad residues characteristic of any cysteine protease (C-H-N), CAP5.5

and CAP5.5V have two major mutations (S-Y-N). Whether they retain any proteolytic activity remains of interest as absence of catalytic residues is not necessarily an indication of enzymatic inactivity (reviewed in Todd *et al.,* 2001). While its domain III has some sequence homology to the C2-L domain of typical domains, little is known about either or their C-terminal domains.

CAP5.5 is one of the few kinetoplastid calpains that are differentially expressed during the life cycle of the parasites. CAP5.5 is expressed only in the procyclic trypomastigotes and is a late-stage marker for the differentiation process from bloodstream forms (Matthews & Gull, 1994). When CAP5.5 transcripts were validated by qRT-PCR, a ratio of 6.0 between procyclic and bloodstream parasites was found (Liu *et al.,* 2010). Similarly, CAP5.5V was also subject to strict stage-specific regulation as the protein is only expressed in bloodstream forms, with a qRT-PCR ratio of 15 between bloodstream and procyclic parasites (Liu *et al.,* 2010).

CAP5.5 is evenly distributed across the sub-pellicular cytoskeleton of procyclic cells, irrespective of the cell cycle position (Hertz-Fowler *et al*., 2001). These observations are consistent with the transcriptome profiling over early G1, late G1, S and M phases of the cell division cycle in *T. brucei*, which show no significant cell-cycle regulation of *CAP5.5* and *CAP5.5V* mRNA (Archer *et al.,* 2011).

RNAi-based functional characterization of both CAP5.5 and CAP5.5V has shown them to be essential for the proper morphogenesis of procyclic and bloodstream trypomastigotes respectively (Olego-Fernandez *et al.,* 2009). In both life cycle stages, depletion of either protein led to a retarded growth rate with an

accumulation of aberrant cells with abnormal kinetoplast/nucleus configurations such as 1K0N (zoids) and 1K2N, which suggested severe perturbations of cytokinesis in addition to the nuclear mispositioning, irregular flagellar pocket biogenesis and the disordered sub-pellicular microtubule corset observed (Olego-Fernandez *et al.,* 2009).

1.2.8 Aims and strategies

In this MSc thesis, I employed a two-pronged approach combining cell biology studies of CAP5.5 and CAP5.5V in *T. brucei* with *in vitro* assays of recombinant calpains expressed in *Escherichia coli* to further characterize CAP5.5 and CAP5.5V.

Are the functions of CAP5.5 and CAP5.5V dependent upon a proteolytic activity, and might these two calpains be druggable targets for the African sleeping sickness? One way to address these questions was by the production of soluble recombinant calpains, determination of proteolytic activity *in vitro* and solving their atomic structures. In chapter 3, I describe my efforts in resolving these issues.

In chapter 4, my aims were two-fold: firstly, to corroborate and further characterize the function of CAP5.5 by over-expressing the full-length protein at two different levels for dominant-negative analyses. Secondly, I investigated the molecular determinants in CAP5.5's sequence such as motifs or domains, which are involved in targeting the dually acylated calpain to the sub-pellicular cytoskeleton. This was achieved by the ectopic expression and subsequent characterization of a series of CAP5.5 deletion mutants.

2. Materials and Methods

2.1 Molecular Biology

2.1.1 Polymerase Chain Reaction

A typical PCR reaction of 50 µl consisted of 0.2 mM of each deoxyribonucleoside triphosphate (dNTPs), 0.2 µM of sense and antisense oligonucleotides (see Table 2.1), 100 ng of trypanosomal genomic DNA and 1 unit of phusion DNA polymerase (NEB; 2 $U.\mu l^{-1}$) in the buffer conditions prescribed by the manufacturers. The primers had restriction enzyme sites incorporated into their sequences for subsequent cloning procedures. *T. brucei brucei* TREU927 genomic DNA was extracted from mid-log phase procyclic cells using Qiamp blood DNA extraction kit (Qiagen). The thermal cycling protocol consisted of an initial denaturation step at 95ºC for 2 minutes immediately followed by 30 cycles of 30 seconds (s) at 95ºC, 30 s at 60ºC and 45 s (per 1 kbp) at 72ºC. The final step consisted of 2 minutes at 72ºC. The reactions took place in a Perkin Elmer GeneAmp 2400 PCR machine. Amplification was assessed by agarose gel electrophoresis and the PCR products purified by either the QIAquick PCR purification kit (Qiagen) or the QIAquick gel extraction kit (Qiagen).

2.1.2 Agarose gel electrophoresis

DNA samples were resolved on a 0.7-1.5% w/v agarose gel that had been melted in TAE buffer (40 mM Tris, 20 mM acetic acid, 10 mM EDTA, 0.5 $\mu g.ml^{-1}$ ethidium bromide) for about an hour at 100 V. DNA bands were visualized using a UV transilluminator.

Table 2.1 List of primers used in this thesis work.

Construct name	Primer sequence (5' to 3'; F=forward; R=reverse)
Recombinant protein expression in *E. coli*	
CAP5.5 FL::His_6 (pET29a)	F: ATAC*ggatcc*ATGGGTTGTGGTGGATCAAAG R: ATA*ctcgag*TTCATCACTGTCAGTTTC
CAP5.5V FL::His_6 (pET29a)	F: ATAC*ggatcc*ATGGGTTGTGGTGGATCAAAG R: ATA*ctcgag*GTCAGTGTCGCTATCATCCGA
CAP5.5V DI-III::His_6 (pET24b)	F: ATA*catatg*GGTGAGGTGACACCG R: ATA*ctcgag*CTTCTCGTCCAAGTACAG
CAP5.5V D II-III::His_6 (pET24b)	F: ATA*catatg*AGCGTGGATGGTAAGAAG R: ATA*ctcgag*CTTCTCGTCCAAGTACAG
CAP5.5V D I::His_6 (pET24b)	F: ATA*catatg*GGTGAGGTGACACCG R: ATA*ctcgag*CATAACTGTCTCCATCTC
CAP5.5V D II::His_6 (pET24b)	F: ATA*catatg*AGCGTGGATGGTAAGAAG R: ATA*ctcgag*CTTCTCGTCCAAGTACAG
CAP5.5V D IV::His_6 (pET24b)	F: ATAcatatgGGTGAGATTCACTTCACC R: ATA*ctcgag*GTCAGTGTCGCTATCATCCGA
Over-expression in *T. brucei*	
CAP5.5::YFP::Ty (pDex577; pDex777)	F: ATA*cttaag*ATGGGTTGTGGTGGATCAAAG R: ATA*actagt*TTCATCACTGTCAGTTTC
GFP::CAP5.5V (pDex777)	F: ATA*tctaga*ATGGGTTGTGGTGGATCAAAG R: ATA*ggatcc*CTAGTCAGTGTCGCTATCATCCGA
GFP::Ty::CAP5.5 (pDex777)	F: ATA*tctaga*ATGGGTTGTGGTGGATCAAAG R: ATA*ggatcc*TTATTCATCACTGTCAGTTTC
CAP5.5 D I-III::YFP::Ty (pDex577)	F: ATA*cttaag*ATGGGTTGTGGTGGATCAAAG R: ATA*actagt*GTCAAGCTTTTCGTCCAATGG
CAP5.5 D I-II::YFP::Ty (pDex577)	F: ATA*cttaag*ATGGGTTGTGGTGGATCAAAG R: ATA*actagt*CTTCTCGTCCAAGTACAGAAC
CAP5.5 DI::YFP::Ty (pDex577)	F: ATA*cttaag*ATGGGTTGTGGTGGATCAAAG R: ATA*actagt*CATGACCGCCTCCACATC
CAP5.5 N37::YFP::Ty (pDex577)	F: ATA*cttaag*ATGGGTTGTGGTGGATCAAAG R: ATA*actagt*CGCTTCCGATGGGACACTTGGCG
CAP5.5 N6::YFP::Ty (pDex577)	F: CTAGTTGATCCACCACAACCCATA R: AGCTTATGGGTTGTGGTGGATCAA
Extending the multiple cloning sites (2303-2264) of pDex577/777	F:AGCTTCTAGAACTTAAGATCGATAGCACTAGTGGAGGTGG AGGTTCAGGAGGTGGAGGTTCTG R:CTAGCAGAACCTCCACCTCCTGAACCTCCACCCCACTAGT GCTATCGATCTTAAGTTCTAGA

Restriction sites are in italics and in a 5' to 3' format. All primers were ordered from Sigma.

2.1.3 Restriction digest and DNA ligation

2 units of restriction enzyme were used for every µg of DNA in manufacturer recommended buffer conditions at 37 ºC for ≤ 2hours. All restriction enzymes used were from New England Biolabs (NEB) or Roche. The digested products were verified by agarose gel electrophoresis prior to purification by either the QIAquick PCR purification kit (Qiagen) or the QIAquick gel extraction kit (Qiagen) according to manufacturer's instructions. Desired constructs were generated by ligating the digested PCR products into the digested vector backbone in a 3:1 insert to vector molar ratio in a 10 µl ligase buffer (NEB) containing 0.5 U of T4 DNA ligase (NEB) for 2 hours at room temperature.

5 µl of ligation products were transformed into competent XL-1 blue strain (Stratagene) of *Escherichia coli* by heat-shocking the cells, which consisted of incubating the ligation mix with the cells for 30 min on ice, 42ºC for 40 s and 8 min on ice. These cells were then plated on LB agar plates containing the appropriate antibiotics (100 $mg.ml^{-1}$ for ampicillin or 50 $mg.ml^{-1}$ for kanamycin) and incubated overnight at 37ºC. Single colonies were verified for proper integration of the desired insert into the vector backbone by colony PCR. Colony PCR differed from the typical PCR reaction described above in the use of primers and DNA. Instead of the genomic trypanosomal DNA, 0.5 µl of the single colony resuspended in 30-50 µl of PBS (Invitrogen) was used. The reverse primers used corresponded to insert amplicons whereas the sense primers corresponded to vector backbones. Following successful verification, the single colonies were incubated overnight in LB buffer (Sigma) containing the appropriate antibiotics at 37ºC for 16 h with shaking at 200 $rpm.min^{-1}$.

2.1.4 DNA isolation

Plasmid DNA was isolated from small scale bacterial cultures (5 ml) according to manufacturer specifications by using Qiaprep spin miniprep kit (Qiagen) or NuceloBond Midi Kit (Machery-Nagel) if using larger scale cultures (~150 ml). All constructs were confirmed by Sanger sequencing (Source Bioscience, Oxford) prior to further analyses.

2.2 Heterologous protein expression and purification

2.2.1 Bacterial expression vectors and host strain

The pET expression system (Novagen) was used to heterologously express CAP5.5 and CAP5.5V. In pET vectors, the genes of interest are under the control of T7 promoter and *lac* operators. Therefore, these plasmids must be transformed into expression hosts bearing the T7 RNA polymerase (λDE3 lysogen) under *lacUV5* control where transcription can be induced exogenously by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG). The expression host used was the competent BL21 (DE3) strain of *E. coli* (Promega).

PCR products corresponding to the full-length calpains were subcloned into the NdeI and XhoI sites of pET29a (Fig. 2.1) and fragments corresponding to the various CAP5.5V domains were subcloned into the NdeI and SalI sites of pET24b (Fig. 2.2). For primer sequences, see Table 2.1. All recombinant proteins expressed were engineered to have a C-terminal hexahistidine tag to facilitate purification. BL21 (DE3) *E. coli* stabilates expressing the rat m-calpain (CAPN2) in pET24d was the kind gift of Peter Davies (Queen's University, Canada).

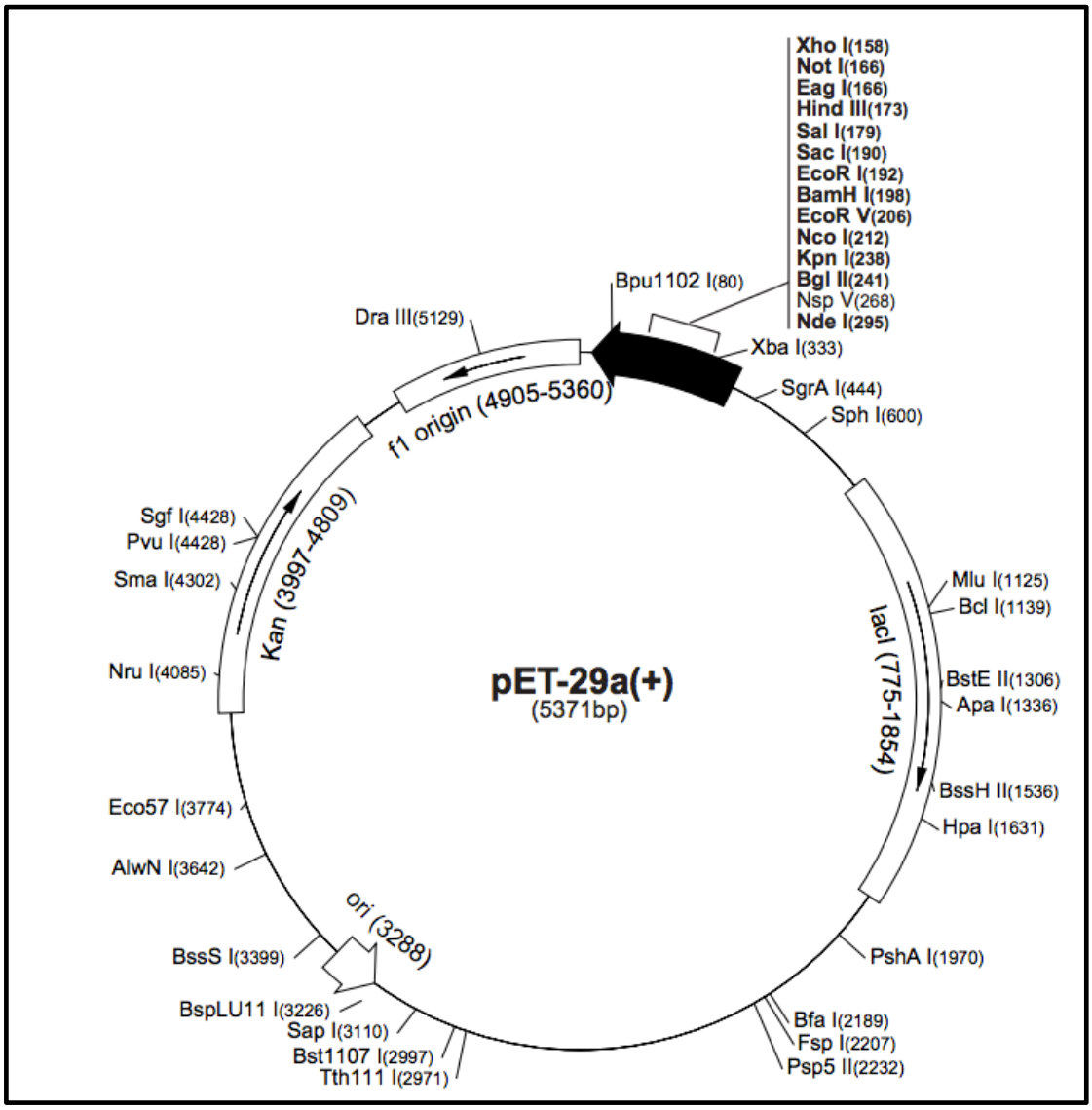


Figure 2.1. Vector map of pET29a used in this thesis. Kan, kanamycin selectable marker; lacI, encodes for a *lac* repressor protein; ori, origin of replication; f1 origin, when exposed to helper phages, produces virions containing single-stranded DNA (not relevant to this study). T7 promoter and the *lac* operator starts 5' of XbaI (333). Image courtesy of Novagen.

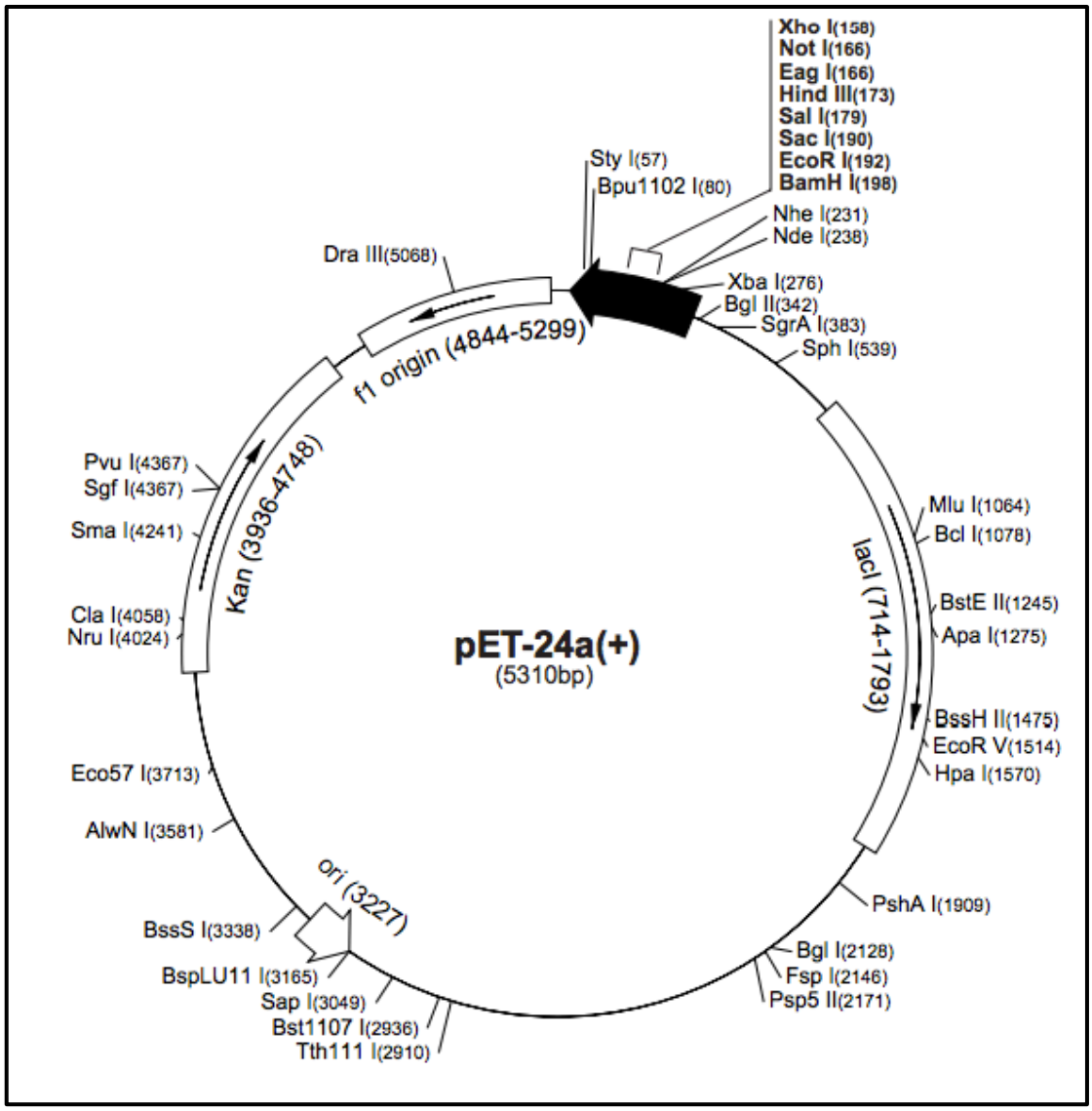


Figure 2.2. Vector map of pET24a. pET24a differs from pET24b as it is 5309 bp compared to 5310 bp of pET24a owing to the additional bp beyond BamHI at 198. Image courtesy of Novagen.

2.2.2 Protein expression conditions

A fresh 5 ml LB culture of BL21 (DE3) cells transformed with the pET vector containing the desired insert was used to inoculate a day culture. This day culture, which consisted of 750 ml of LB medium with 50 mg/ml of kanamycin and 0.8% w/v glucose, was incubated in a shaking incubator at 37ºC until OD_{600} 0.6 > < 1.0 was reached. Upon which, the culture was induced at 22ºC for 16 h with 0.5-1.0 mM IPTG. The glucose was added as a catabolite repression technique to repress basal expression of target genes (Pan & Malcolm, 2000).

Cells were harvested by centrifugation (5100 x *g*, 30 min, 4ºC), resuspended in ice-cold lysis buffer (25 mM Tris (pH7.6), 5 mM imidazole, 1 mM EDTA, 5% v/v glycerol, 4 mM BME, Roche protease inhibitor cocktail) and homogenized at 15, 000 psi (5 times) using an Emulsiflex-C5 homogenizer (GC Technologies). The homogenate was fractionated by centrifugation (21, 000 x *g*, 45 min, 4ºC) into insoluble (pellet) and soluble fractions (supernatant).

2.2.3 Nickel affinity chromatography

The proteus IMAC midi spin column kit (AbD Serotec) was used for nickel affinity based purification of the hexahistidine tagged recombinant proteins. Prior to nickel affinity chromatography, 20 mM $MgCl_2$ (final concentration) was added to the soluble fraction in order to chelate EDTA. The soluble fraction was applied onto a 10 ml column containing Ni^{2+} immobilized resin, washed in buffer (25 mM Tris (pH 7.6), 500 mM NaCl, 5% glycerol, 4 mM BME) containing 35 mM imidazole and

eluted with 300 mM imidazole into a collection fluid containing 0.1 M EDTA, 20 mM Tris (pH7.6), and 10 mM BME.

Eluted fractions were dialyzed overnight at 4ºC in a 10, 000 dalton protein molecular weight cut-off "Slide-a-lyzer" dialysis cassette (Pierce) against 1 l dialysis buffer (20 mM Tris (pH7.6), 20 mM NaCl, 2% glycerol, 10 mM BME).

2.2.4 Size exclusion chromatography and MALLS

The dialyzed proteins were further purified by size exclusion chromatography using HiLoad 16/60 Superdex 200 pg column (GE Healthcare) that had been pre-equilibrated with 20 mM Tris (pH7.6), 20 mM NaCl, 2% glycerol, 10 mM BME at 2 $ml.min^{-1}$. Eluted fractions were collected and concentrated to 2-5 mM using an Amicon centrifugal filter unit (Millipore).

For multi-angle laser light scattering (MALLS), protein samples were processed similarly to size exclusion chromatography except the S200 16/60 column was followed in-line by a Dawn Heleos-II light scattering detector and an Optilab-Rex refractive index monitor (Wyatt Technologies). Molecular mass calculations were performed using ASTRA 5.3.4.14 (Wyatt Technologies) assuming a refractive index increment (dn/dc) value of 0.186 $ml.g^{-1}$. All MALLS analyses were performed by Steven Johnson.

2.2.5 Crystallization screening

Several 96 well commercial sparse matrix screens were tested for CAP5.5V dII-His_6 crystallization. They included PACT premier HT-96, JCSG-plus, Structure Screen I & II, Morpheus, ProPlex (all from Molecular Dimensions). Taken together, these screens tested hundreds of different conditions incorporating salts, polyethylene glycol, neutralized organic acids or organic precipitants across a pH range from 4.0 to 10.5, including a diverse range of salt additives. Each crystallant solution (300 nl) in the 96 well plates were tested at 21ºC by sitting drop vapour diffusion using two different protein concentrations of 50% v/v and 75% v/v of CAP5.5V dII at ~14.4 $mg.ml^{-1}$. The drops were set up with an OryX nano crystallization robot (Douglas Instruments).

2.2.6 Protease activity assays

Calpain activity of nickel affinity purified proteins was measured using the Calpain-Glo™ protease assay kit (Promega). This chemiluminescent reaction was initiated by adding the nickel eluates to the succinyl proluminescent substrate, suc-LLVY- aminoluciferin, diluted 1:10 in the assay buffer containing the luciferin detection reagent and the calpain-Glo™ buffer. Activity was assessed by varying the calcium concentration (2mM or no $CaCl_2$) and/or in the presence of protease inhibitors cocktail (Roche). The solution was gently mixed at room temperature for 30 minutes before loading them onto 96-well white-walled plates for measurement in a luminometer (Wallace 1420 Multilabel Counter, PerkinElmer).

Another assay used to assess calpain activity was done by incubating equal concentrations (10 μg) of purified calpains and commercial-grade substrates such as casein (Sigma) and bovine serum albumin (BSA; New England Biolabs) in 50 μl of activity buffer (20 mM Tris (pH 7.6), 20 mM NaCl, 10 mM BME, 2% glycerol, 10 mM $CaCl_2$) at a 37ºC water bath incubator for 1 h. As a positive control, rat CAPN2 and trypsin (Roche) was used. Proteolytic degradation, if any, was visualized on a Coomassie-stained SDS-PAGE gel.

2.2.7 Polyacrylamide gel electrophoresis

Several different kinds of polyacrylamide gels were run and protein samples were prepared accordingly. 10% and 4-12% Criterion XT bis-tris polyacrylamide gels, which contain no SDS, were used for all analyses (BioRad). Apparent molecular weight was determined by using Precision Plus protein markers (BioRad). Visualization of proteins on these gels was carried out using either Coomassie blue stain (2.5 g coomassie brilliant blue R250 in 45% methanol, 10% acetic acid) or Sypro Ruby stain (BioRad) overnight at room temperature. Sypro-Ruby stained gels were visualized by a UV transilluminator and Coomassie stained gels were destained in a solution consisting of 7% acetic acid and 45% methanol.

The standard **SDS-PAGE**, for both bacterial and trypanosomal cell lysates, was run using samples solubilised in boiling Laemmli sample buffer (62.5 mM Tris (pH 6.8); 40% glycerol; 0.01% bromophenol blue; 2% SDS, 700 mM BME) at 90ºC for 5 min. Gels were run in XT MOPs running buffer (BioRad), which contain 0.15% w/v SDS at a current of 40 milliamperes (mA) for 1 h. Variations on this standard

SDS PAGE that were used in this thesis include using unboiled protein samples or using protein samples solubilised in Laemmli buffer without any reducing agents such as ß-mercaptoethanol (BME; Sigma).

Native PAGE consisted of proteins solubilised in native sample buffer (62.5 mM Tris-HCl, pH 6.8; 40% glycerol; 0.01% bromophenol blue), not boiled, and run in a running buffer (50 mM Tris (pH7.7), 50 mM MOPS, 1 mM EDTA) at 40 mA for ~3 h.

2.2.8 Optimization trials

While the above represents the fully optimized protocol, several optimization trials were performed where the protein samples were incubated in different buffers at 37ºC for 16 h, and aggregation assessed by gel or chromatographic analyses. The following variables were changed: buffer pH (5.1- 7.0); concentration (5, 10, 30 mM) and identity of reducing agents (all from Sigma) such as ß-mercaptoethanol (BME), dithiothreitol (DTT) and tris (2-carboxyethyl) phosphine (TCEP); detergents (all from Sigma) such as octyl glucopyranoside, triton X-100, triton X-114, IGEPAL (at concentrations below their respective critical micelle concentrations) and many others. pH 7.0 buffers consisted of 75 mM MOPS, 110 mM NaCl and pH 5.1 buffers were made of 75 mM NaCitrate, 110 mM NaCl. These buffers either contained 0, 5 mM or 30 mM of different reducing agents such as BME, DTT or TCEP. The detergents were used in buffers of various pH at 0.3% v/v of either octyl glucopyranoside, triton X-100, triton X-114, or IGEPAL.

In order to irreversibly acetylate free cysteines on CAP5.5V, the nickel eluates were incubated in a buffer containing 50 mM ammonium bisulfate (pH8.0), 90 mM iodoacetamide for an hour at room temperature prior to dialysis against the gel filtration buffer (25 mM Tris (pH7.6), 145 mM NaCl, 2% glycerol, 10 mM BME).

2.3 *T. brucei* cell lines and cell culture

2.3.1 Expression vectors

The promoter-less polycistronic transcription in trypanosomatids has been exploited to control gene expression exogenously. For instance, use of inducible expression plasmids in cell lines that stably express both T7 RNA polymerase and tetracycline repressor (TetR) protein have allowed for inducible RNA interference and inducible transgene expression experiments (Poon *et al.,* 2012; Wirtz & Clayton, 1995). Here, transgene expression can be controlled in an inducer-dependent manner as the chromosomally integrated gene of interest is under the control of a T7 RNA polymerase promoter and tetracycline operators (Poon *et al.,* 2012).

The two ectopic expression vectors used in this study were pDex577-YFP and pDex777-GFP (Fig. 2.3; Poon *et al.,* 2012). The plasmids integrate into the mini chromosomes and encode YFP/GFP genes under the control of a T7 RNA polymerase promoter (Poon *et al.,* 2012). The primary difference between the two vectors is the 3' intergenic regions of the induced gene; pDex577 has an aldolase 3' UTR and pDex777 has paraflagellar rod 2 (PFR-2) 3' UTR immediately downstream of the YFP/GFP-coding sequences. CAP5.5 and CAP5.5V with a N-terminal GFP were

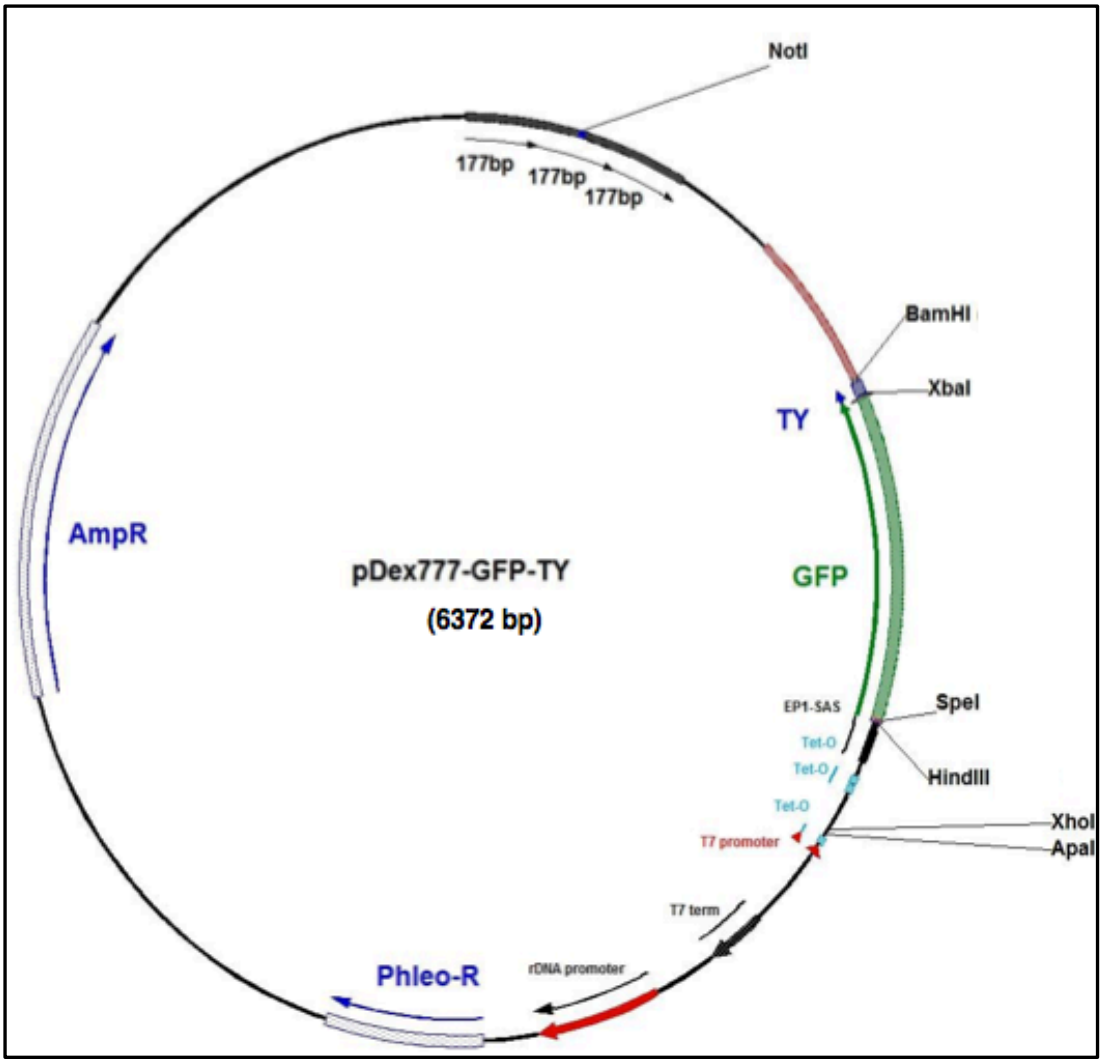


Figure 2.3. Vector map of pDex777-GFP-Ty used for ectopic expression of CAP5.5. The sole difference in the vector map of the pDex577 used is that the 3'UTR paraflagellar rod 2 (PFR-2) has been replaced by aldolase intergenic sequences**.**

engineered by subcloning amplicons corresponding to the full length protein into the XbaI and BamHI sites of pDex777.

Owing to the limited choices of restriction enzyme sites (HindIII and SpeI only) for subcloning DNA fragments into pDex vectors to engineer C-terminal GFP tagged proteins and the finding of several native HindIII sites in CAP5.5/CAP5.5V, an extended multiple cloning region was subcloned in frame to the HindIII and SpeI sites of pDex577 and pDex777. The 63 bp DNA fragment containing the new multiple cloning sites was ordered from Sigma and used as annealed, reverse complementary primers. This fragment also encoded the GS_{10} linker region used to separate the protein of interest from GFP. DNA amplicons encoding the various CAP5.5 domains and motifs were then subcloned into AflII and SpeI sites. For the sequences of the oligonucleotides used, please refer to Table 2.1.

2.3.2 Trypanosome cell culture

Procyclic *T. brucei brucei* SmOx P927 cell lines were cultured at 28ºC in SDM-79 medium supplemented with 10% v/v heat-inactivated fetal bovine serum (Gibco) in 25cm^2 tissue culture flasks (Brun & Jenni, 1977). Bloodstream SmOx B427 forms were cultured at 37ºC in HMI-9 medium supplemented with 15% v/v heat-inactivated fetal bovine serum (Gibco) in a vented 25cm^2 tissue culture flask with 5% carbon dioxide (Hirumi & Hirumi, 1989).

SmOx P927 and B427 are TREU927 and Lister 427 cells respectively that express T7 RNA polymerase and tetracycline repressor (Poon, S. K. *et al.,* 2012).

These cells were maintained in the logarithmic phases of their growth by repeatedly passaging the cultures to a cell density of 1 x 10^5 to 2 x 10^7 cells.ml^{-1} for procyclic cells and a maximum density of 1 x 10^6 cells.ml^{-1} for bloodstream forms. Cell density was measured using a model TT Casy counter (Schärfe Systems).

The culture media of cell lines were maintained under selection with the appropriate drugs: phleomycin (5 μg.ml^{-1} for procyclic and 2.5 μg.ml^{-1} for bloodstream cells); puromycin (1 μg.ml^{-1} for procyclic and 0.2 μg.ml^{-1} for bloodstream cells). Ectopic expression was induced by the addition of tetracycline at a final concentration of 1 μg.ml^{-1}.

2.3.3 Transfection of trypanosomes

For each transfection, logarithmic-phase trypanosomes with a density of 1 x 10^7 cell.ml^{-1} were harvested by centrifugation (800 x *g*, 10 min, room temperature), washed once in cytomix buffer (25 mM HEPES (pH7.6), 120 mM KCl, 0.15 mM $CaCl_2$, 2 mM EGTA, 5 mM $MgCl_2$, 0.5% w/v glucose (dextrose), 100 μg.ml^{-1} BSA, 1 mM hypoxanthine, 10 mM K_2HPO_4/KH_2PO_4) at room temperature and resuspended in cytomix. This cell suspension was mixed with 10 μg of NotI linearized plasmid (in EB buffer, NEB) in a 4mm electroporation cuvette that was then electroporated with 2 consecutive 100 μs square pulses (1 if using bloodstream cells) at 1.7 kV. An Electro square porator ECM830 (BTX) was used. After transfection, parasites were transferred to 10 ml medium (25 ml for bloodstream cells) and cultivated without selection for 16 h.

Selection was subsequently applied to the bloodstream cells by diluting the culture 1:20 in HMI-9 medium containing the appropriate antibiotics and 1.5 ml of culture was seeded onto each well of the 24-well microtiter plates.

Procyclic cells were diluted 1:20 and 1:200 in SDM-79 and seeded onto 96-well microtiter plates containing the appropriate antibiotics with 20% v/v conditioned medium. Conditioned medium is the filtered supernatant of a procyclic culture grown to 5 x 10^7 cells.ml^{-1} (lag phase) that has been subsequently harvested (800 x *g*, 10 min, room temperature) and filter-sterilized.

For all experiments, only clones obtained by limiting dilution were analysed unless otherwise noted.

2.4 Protein biochemistry (*T. brucei*)

2.4.1 Preparation of protein samples

Trypanosomes in exponential growth phase were harvested (800 x *g*, 10 min, room temperature), washed in PBS (Invitrogen) and the pellet solubilized in boiling Laemmli sample buffer (62.5 mM Tris (pH 6.8); 40% glycerol; 0.01% bromophenol blue; 2% SDS, 700 mM BME) at 90ºC for 5 min. 3.5 x 10^6 cells.ml^{-1} were loaded per lane for all the gels.

2.4.2 Western blot and immunodetection

SDS-PAGE was done using the same parameters as described with the bacterial cell lysates (Chapter 2.2.7). Following SDS-PAGE, proteins were

transferred onto a nitrocellulose membrane (Whatman) at a constant current of 120 mA for 16 hours in 5 l of transfer buffer (24 mM Tris, 192 mM glycine, 20% MeOH, 0.01% SDS). Success of the transfer was visualized by staining the membrane with Ponceau S stain (0.1% w/v Ponceau S stain in 5% v/v acetic acid).

Membranes were blocked in 5% milk in TBST (10 mM Tris (pH 7.4), 140 mM NaCl, 0.05% Tween-20) for an hour. The membranes were subsequently incubated in 2.5% milk containing TBST buffers containing primary antibody for an hour and then, buffers containing the secondary antibody conjugated with horseradish peroxidase for another hour. Between every step of blocking and antibody application, the membranes were washed for 5 min x 4 with TBST. Each step took place at room temperature with gentle shaking. The membranes were treated with enhanced chemoluminescence reagent (Perkin Elmer), exposed to Kodak films and developed in a Kodak M35-M X-OMAT processor.

The primary antibodies (and their dilutions) used for western blots were as follows: Anti-CAP5.5 mouse IgG monoclonal antibody (1:20 dilution; Ersfeld *et al.,* 2001) and anti-GFP mouse IgG monoclonal antibody (1:500 dilution; 11814460001, Roche).

2.5 Light Microscopy

Unless otherwise stated, all microscopy slides were examined on a Zeiss Axioplan 2 microscope 40x 0.8 N/A oil immersion lens and captured on a CCD camera controlled by Micro-Manager 1.4 (Edelstein *et al.,* 2010). All images were processed using ImageJ software (Abramoff, 2004). Cellular dimensions were also

measured using ImageJ and statistical analyses performed using Prism 6 (GraphPad software).

2.5.1 Native fluorescence microscopy

For the analysis of native GFP/YFP fluorescence in live cells, procyclic trypanosomes were washed in PBS (Invitrogen) containing 1 $\mu g.ml^{-1}$ Hoechst 33324 and subsequently visualized on a microscope. Otherwise, cells were fixed in 2% paraformaldehyde in PBS for 15 minutes followed by three immersion washes in PBS (Invitrogen) before being mounted in glycerol containing 1% w/v 1,4-diazabicyclo [2.2.2] octane (DABCO), 10% v/v 50 mM sodium phosphate (pH 8.0) and 1 $\mu g.ml^{-1}$ of DAPI.

2.5.2 Immunofluorescence microscopy

Trypanosome cells were settled onto glass slides (glutaraldehyde-derivatized slides for bloodstream cells), washed in PBS (Invitrogen), and fixed in 2% paraformaldehyde for 15 minutes. Following permeabilization with 0.1% Triton X-100, 0.1 M glycine in PBS for 10 minutes, cells were labelled with the primary antibody for 1 hour at room temperature and the secondary antibody for also 1 hour at room temperature. Between antibody applications, 4 x 5 min PBS (Invitrogen) immersion washes were done. Samples were mounted in glycerol containing 1% w/v DABCO, 10% v/v 50 mM sodium phosphate (pH 8.0) and 1 $\mu g\ ml^{-1}$ of DAPI.

The primary antibodies used were anti-GFP (rabbit; polyclonal) diluted 1:200 in blocking medium (2% v/v heat-inactivated fetal bovine serum (Gibco), 0.1% v/v BSA, 0.1% v/v sodium azide) and anti-tyrosinated α-tubulin, YL1/2 (rat IgG) used

1:200 in blocking medium (Kilmartin *et al.*, 1982). The secondary antibodies used were Alexa Fluor 594 goat-anti rabbit IgG (A11-037, Invitrogen) and rhodamine isothiocyanate goat anti-rat IgG (112-026-003, Jackson ImmunoResearch). Both secondary antibodies were used 1:500 in blocking medium.

For cytoskeletons, PBS-washed cells on glass slides were treated with 1% Nonidet P-40 in PEME (100 mM PIPES, 2 mM EGTA, 0.1 mM EDTA, 1 mM MgSO4, pH 6.9) for 2 minutes before fixing them in ice-cold methanol at -20°C for 30 minutes. Detergent extraction was verified by light microscopy prior to proceeding. Following rehydration by immersing the slides in PBS for 4 x 5 minutes, the cells were labelled with antibodies in the same way as whole cells.

2.6 Bioinformatics

2.6.1 Amino acid alignment

Multiple sequence alignments were generated using the MAFFT method version 7 L-INS-I (http://mafft.cbrc.jp/alignment/server/). After manual editing, alignments were displayed and shaded with Boxshade at http://www.ch.embnet.org and further annotated in Microsoft Word. PFAM domains were identified on http://pfam.sanger.ac.uk/ .

2.6.2 Homology models

Homology models were generated by threading protein sequence through Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) and FFAS servers (http://ffas.ljcrf.edu/ffas-cgi/cgi/document.pl?ses). These programs build

structural models using the known structures of homologous domains as templates, integrating both sequence and structural -secondary and tertiary- information (Kelley *et al*., 2009; Zhang *et al.*, 2009). The library of known structures are compiled from the Structural Classification of Proteins (SCOP) database and augmented with the Protein Data Bank (PDB) deposits. They contain structures determined by both X-ray diffraction and nuclear magnetic resonance (NMR) spectroscopy. The models were processed in PyMol 1.5.0 (Schrödinger).

3. Expression, purification and biochemical analyses of recombinant calpains

The MALLS work presented in figures 3.6 and 3.11 was contributed by Steven Johnson (University of Oxford).

3.1 Introduction

Systematic analyses of kinetoplastid genomes revealed a large and diverse family of atypical calpains, so called because they lack the PEF domain, in *Trypanosoma brucei, Trypanosoma cruzi* and *Leishmania major* (Ersfeld *et al.,* 2005). Additionally, the majority of these calpain-like proteins deviate from typical calpains for they lack the canonical catalytic triad residues (C-H-N). Cytoskeleton-associated protein 5.5 (CAP5.5) and its paralogous variant, CAP5.5V, are two such atypical calpains that have been shown to play essential roles in the proper morphogenesis of the procyclic and bloodstream forms of *T. brucei* respectively (Olego-Fernandez *et al.,* 2009). However, some key issues remain unresolved: Are the functions of CAP5.5 and CAP5.5V dependent upon a proteolytic activity, and might these two calpains be druggable targets for the African sleeping sickness? One way to address these questions is by the production of soluble recombinant calpains, determination of proteolytic activity *in vitro* and solving their atomic structures. In this chapter, I describe my efforts in resolving these issues.

3.2 Bioinformatic analyses of *T. brucei* CAP5.5 and CAP5.5V

CAP5.5 is encoded by Tb927.04.3950 and CAP5.5V by Tb927.08.8330. CAP5.5 is 853 amino acids long with a predicted molecular weight of 94.6 kDa and an isoelectric point of 4.15. Compared to CAP5.5, CAP5.5V is a larger protein with 888 amino acids and a predicted molecular weight of 98.4 kDa but with a similar isoelectric point of 4.27 (Table 3.1). Both calpains contain PFAM peptidase C2 and

DUF1935 domains, which correspond to domain II and domain I respectively. The sequence identity between CAP5.5 and CAP5.5V for the first 720 amino acids of CAP5.5 is 84% and 75% across their full lengths, differing primarily in their C-terminal domain. As mentioned in the introduction, most of the kinetoplastid calpains were identified based on the significant homology between their domain IIs and the protease core domains of canonical calpains (Ersfeld *et al.,* 2005). The amino acid alignment of CAP5.5 and CAP5.5V with human calpains 1 (CAPN1) and 2 (CAPN2) re-emphasizes the homology between the domain IIs (Fig 3.1). The alignment between CAP5.5's domain II and the catalytic domains of human CAPN1 and CAPN2 showed that 23% of amino acids exhibited identity, with 41% of the residues being similar in their physicochemical properties. The alignment also highlighted CAP5.5 and CAP5.5V's degenerate catalytic triad residues. The calcium binding residues in domain II of typical calpains are acidic in nature to facilitate conformational activation of typical calpains by forming salt bridges. In CAP5.5 and CAP5.5V, only 3 out of the 5 predicted calcium-binding residues are acidic in nature, with a key residue (E302 in CAPN2) mutated to a neutral, non-polar residue (P483 in CAP5.5V). Mutational analyses of CAPN2 have shown E302R mutants have severely impaired calpain activation and activity as such a mutation eliminated a key salt bridge (Moldoveanu *et al.,* 2004b). Therefore, calcium regulation for CAP5.5/CAP5.5V is unlikely.

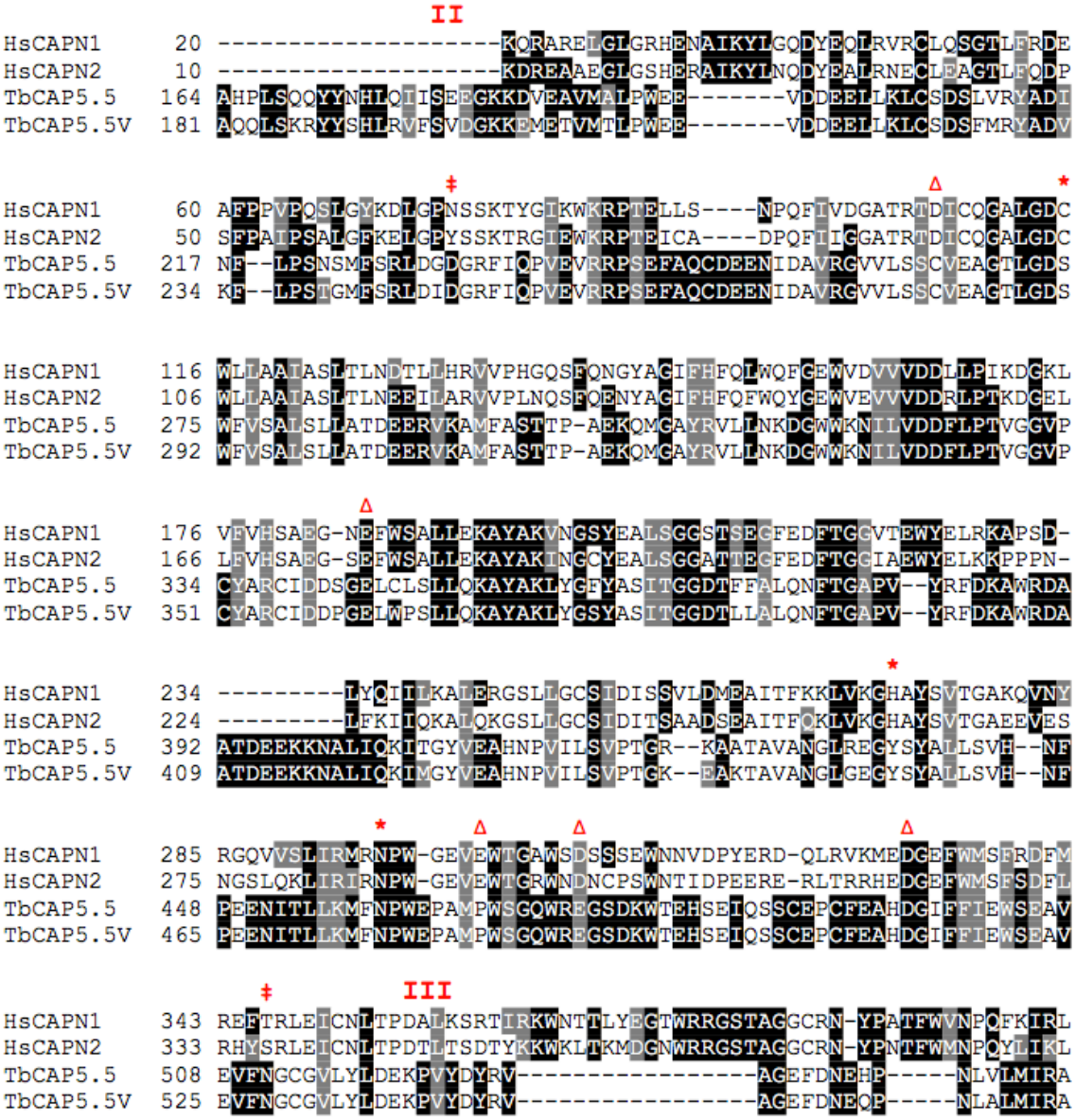


Figure 3.1. Amino acid alignment of the catalytic domains of human CAPN1 and CAPN2 with the corresponding domains of CAP5.5 and CAP5.5V. Alignment was generated using the MAFFT method version 7 L-INS-I (http://mafft.cbrc.jp/alignment/server/). Identical residues are in black and those that share physicochemical properties are in gray; box designations (II and III) are according to domain parsing based upon homology modelling. The predicted "catalytic" triad residues are denoted with a red asterisk (*); the calcium-binding acidic residues with a red triangle (Δ), and the boundaries of PFAM peptidase C2 domain are denoted with a red diesis (‡). Abbreviations (with NCBI accession numbers in parentheses) are as follows: Calpain 1 *Homo sapiens* HsCAPN1 (NP_005177); Calpain 2 *Homo sapiens* HsCAPN2 (NP_001739); CAP5.5 *T. brucei* (AAG48626); CAP5.5V *T. brucei* (AAZ13590.1).

	*Hs*CAPN1	*Hs*CAPN2	*Tb*CAP5.5	*Tb*CAP5.5V
NCBI Accession ID	NP_005177	NP_001739	AAG48626	AAZ13590.1
Molecular Weight (kDa)	81.8	79.9	94.6	98.4
Isoelectric point (pI)	5.5	4.9	4.2	4.3
PFAM domains (e-value)	Peptidase C2 (1.7e-150) Calpain domain III (3.2e-58) EF Hand (6.8e-08)	Peptidase C2 (2.1e-152) Calpain domain III (1.6e-58) EF Hand (1.7e-07)	Peptidase C2 (1.4e-36) DUF1935 (7.6e-25)	Peptidase C2 (5.1e-39) DUF1935 (1.3e-23)
Total number of cysteines	11	15	21	20

Table 3.1. Biochemical properties of CAP5.5 and CAP5.5V. Selected biochemical properties of CAP5.5, CAP5.5V, CAPN1 and CAPN2 are highlighted here. The isoelectric point (pI) estimates assume that all residues have pKa values that are equivalent to the isolated residues. The PFAM domains were identified on http://pfam.sanger.ac.uk/. Peptidase C2 refers to the catalytic domain II of typical calpains; EF hand to the calmodulin-like calcium-binding domain IV of typical calpains and DUF1935 is a domain of unknown function that is homologous to domain I^K of kinetoplastid calpains. The e-value (expectation value) is the number of hits that would be expected to have a score equal to or better than this value by chance alone. While these cytosolic proteins are unable to form disulfide bonds, once heterologous cells have been lysed and the proteins are in an artificial condition, the probability of disulfide bond formation exists. Therefore, the number of cysteines in each protein is presented here to highlight the possibility of disulfide bonds formation when working with recombinant cytosolic proteins in solution.

Such sequence homology was used in conjunction with secondary structure predictions and homology models to delineate the structural domain boundaries of CAP5.5V in order to create a series of constructs for crystallographic analyses. Only the domains of CAP5.5V were heterologously expressed.

Homology models were generated by threading CAP5.5V protein sequence through Phyre2 and FFAS servers. These programs build structural models using the known structures of homologous domains as templates, integrating both sequence and structural -secondary and tertiary- information (Kelley *et al.,* 2009; Zhang *et al.,* 2009). The library of known structures are compiled from the Structural Classification of Proteins (SCOP) database and augmented with the Protein Data Bank (PDB) deposits. They contain structures determined by both X-ray diffraction and nuclear magnetic resonance (NMR) spectroscopy. The model for CAP5.5V domain I was based on the structure of *L. major* small myristoylated protein-1 (SMP-1; protein database entry 2FEO), which shares a 36% sequence identity (Fig. 3.2A). The highly disordered region encompassing the 92 amino terminal residues (N92), which includes the fatty acid acylation motif and the poly-proline stretch, were excluded from heterologous expression constructs as it was hypothesized to impede crystal packing.

Homology models for CAP5.5V domain II (26% sequence identity) and III (16% sequence identity) were derived using the structure of rat CAPN1 (PDB: 1QXP) as template (Fig. 3.2A and 3.2B). The extended amino terminal α-helix (50 aa) preceding domain II was included in expression constructs as a possible electrostatic interaction was hypothesized between E204 and W333-W334-K335. Additionally,

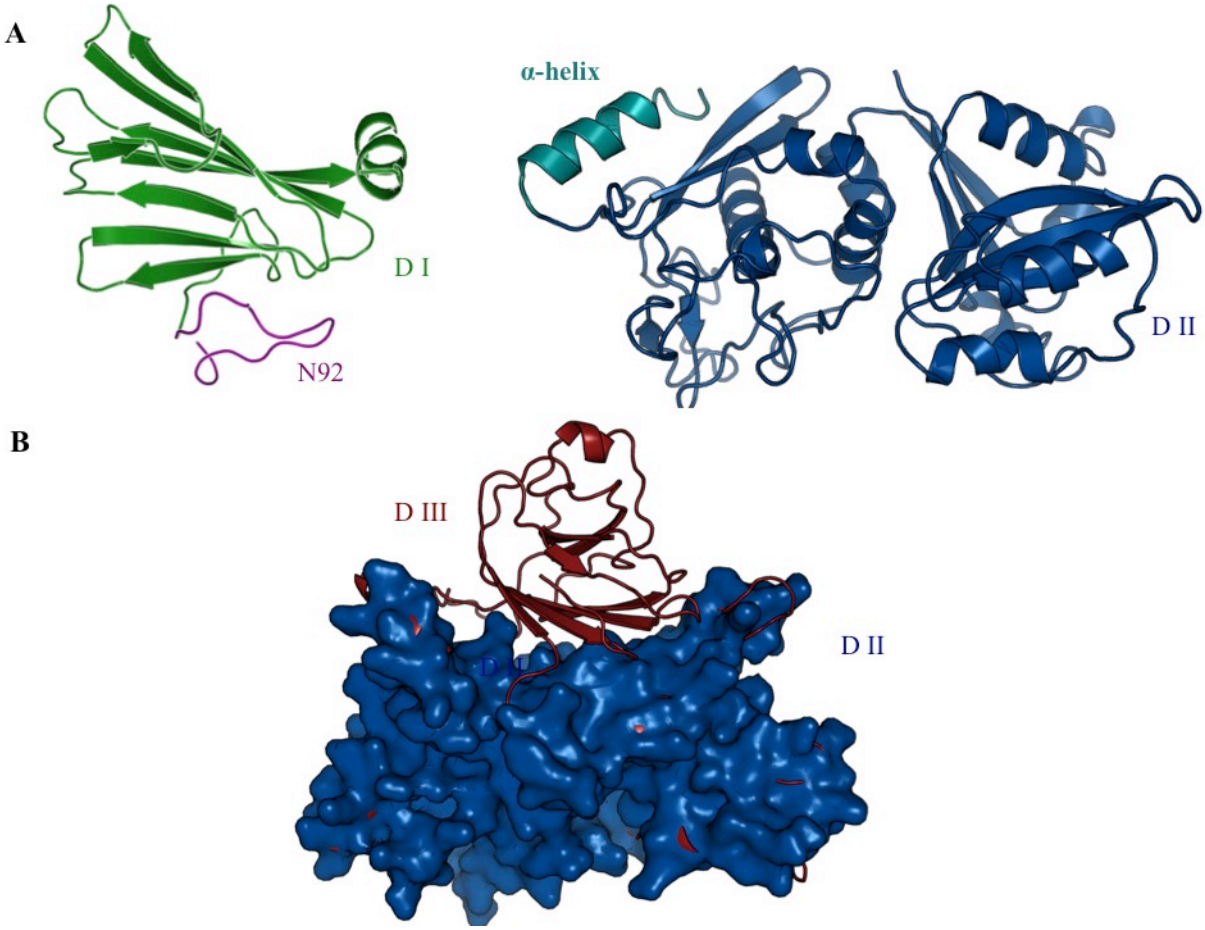


Figure 3.2. Homology modelling and domain parsing of CAP5.5V. A. Predicted cartoon model of CAP5.5V domain I (D I) including the first 92 N-terminal amino acids (N92) is based on the NMR structure of *L. major* SMP-1 (PDB: 2FEO) and domain II (D II) based on rat CAPN1 (PDB: 1QXP). The α-helix (α-helix) preceding domain II was hypothesized to be involved in a stabilizing electrostatic interaction and was included in all constructs expressing domain II. **B.** Homology modelling highlights the interaction between the predicted domain III (D III) structure (PDB: 1QXP) shown as ribbons and the surface diagram of domain II. Domains are shown in similar colours and all structures are oriented with the N-terminus to the left. The homology models were generated using Phyre2.

two constructs were designed to preserve the predicted hydrophobic interaction between domain II and domain III by designing domain I-II-III and domain II-II expression constructs (Fig 3.2B). No protein structure was predicted for domain IV from the analyses due to a lack of homologous domains. Based on this analysis, CAP5.5V domain boundaries were predicted and a series of constructs were designed (Fig 3.3). In addition to the full-length CAP5.5 (1-853) and CAP5.5V (1-888), CAP5.5V domain I (92-207), domain II (196-538), domain I-III (92-676), domain II-III (196-676), and domain IV (675-888) were engineered into pET expression plasmids fused to C-terminal hexahistidine tags.

3.3 Soluble CAP5.5 and CAP5.5V expression

In order to determine the expression profiles of CAP5.5 (1-853) and CAP5.5V (1-888), protein expression was induced at 24°C by 0.5 mM IPTG overnight, homogenized, centrifuged to fractionate the insoluble and soluble materials, and the soluble fraction (supernatant) was subjected to nickel affinity chromatography. The cell lysates from the initial expression trials were visualized by SDS-PAGE and showed soluble recombinant proteins that were later further isolated by the affinity column (Fig. 3.4A). I was also able to successfully induce the expression of soluble rat m-calpain (CAPN2), which has been demonstrated to be an active protease, to be used as positive controls for the calpain activity assays (Larsen *et al*., 2004).

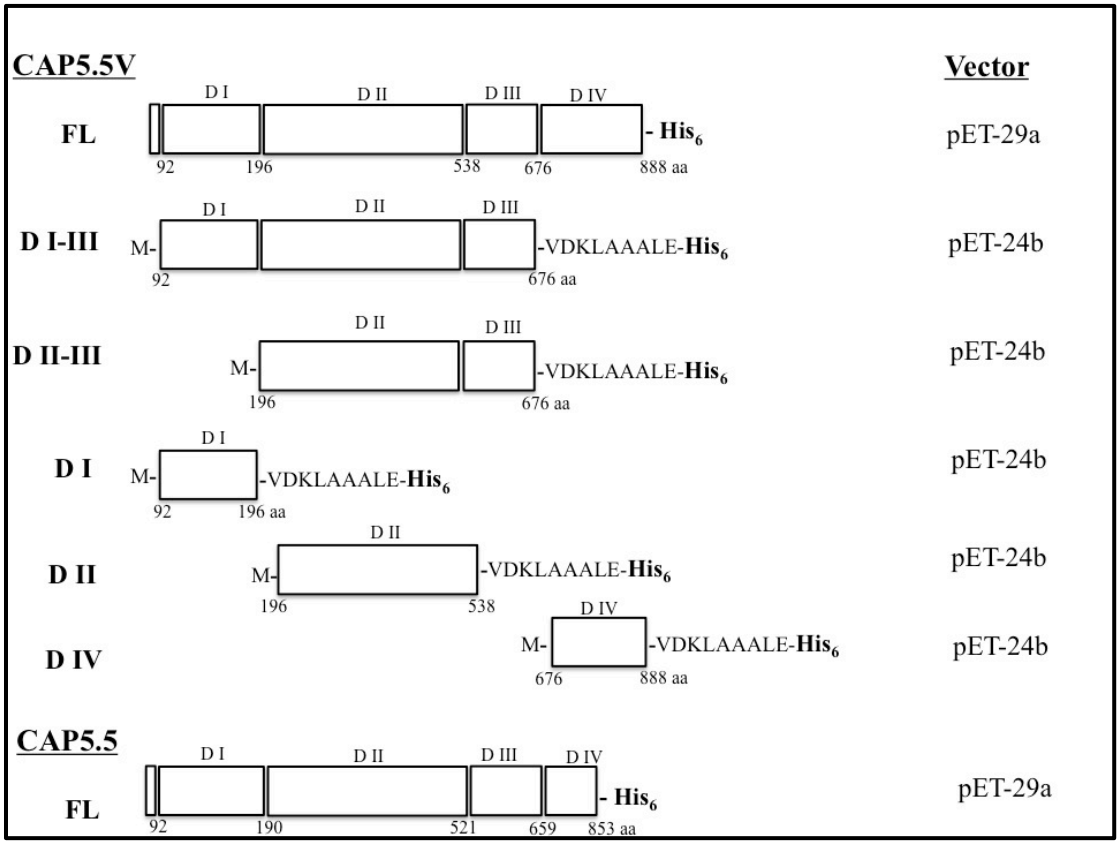


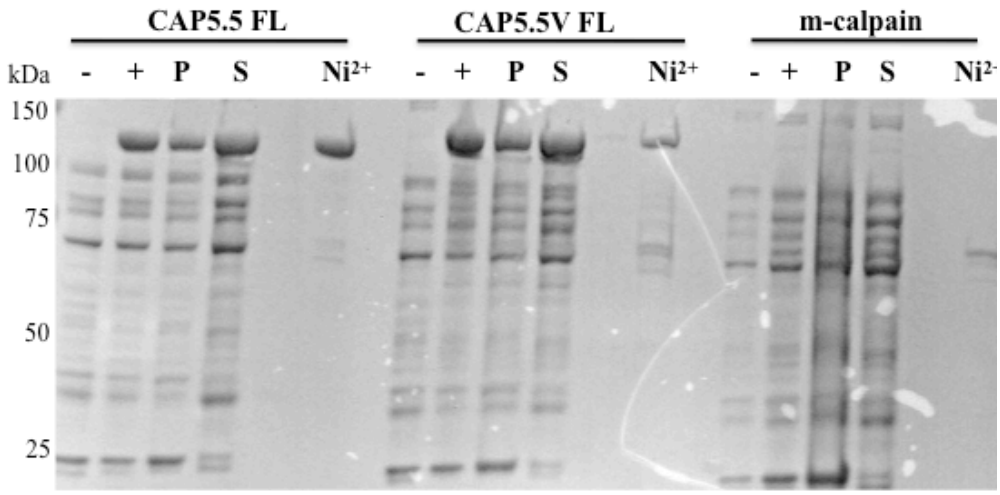
Figure 3.3. Schematic representation of the expression constructs made in this study. The coordinates of the domain boundaries of CAP5.5 and CAP5.5V are shown as well as those of the series of constructs expressed in *E. coli* for biochemical analyses. Expression in both pET-29a and pET-24b is under the control of a T7 promoter and utilizes T7 terminator sites. The two constructs differ only in their multiple cloning regions. All constructs were expressed with C-terminal hexahistidine (His_6) tags in order to aid in protein purification. Four constructs were expressed with a linker region (VDKLAAALE) fused to the hexahistidine tag (as expressed in Elce *et al.,* 1995).

To investigate CAP5.5 and CAP5.5V's protease activity, a commercial, luminescent assay was used that provides a succinyl, proluminescent calpain substrate, Suc-LLVY-aminoluciferin. Upon cleavage of the substrate, a luminescent signal is produced by the luciferase reaction that is detectable by a luminometer, where the signal is proportional to the amount of calpain activity present (Fig. 3.4B). The nickel affinity eluates were incubated in the presence of 2 mM free Ca^{2+}, 0 Ca^{2+}, or with 2 mM free Ca^{2+} and a cocktail of protease inhibitors. These conditions were chosen because calcium binding is required for the full enzymatic activity of canonical calpains, which should be suppressed upon the addition of inhibitors (Moldoveanu *et al.,* 2002).

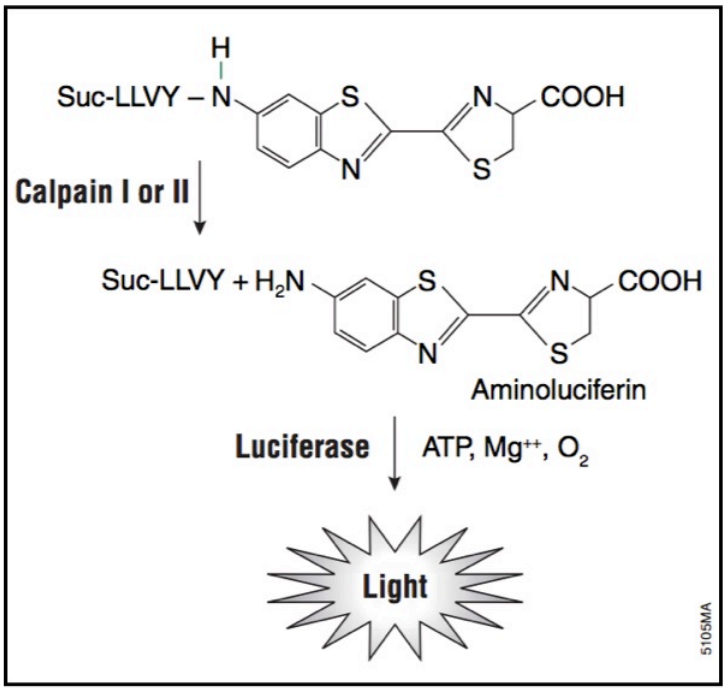
For this preliminary assay, the eluates were used at the same concentrations despite not being purified to homogeneity. Little to no signal was observed in the absence of Ca^{2+} or the presence of Ca^{2+} regardless of the presence of protease inhibitors with recombinant CAP5.5 and CAP5.5V (Fig. 3.4C). The difference in the relative luminescence units between the assay buffer and CAP5.5/V may be better controlled for in future experiments with a purified protein with no known protease activity such as GFP/YFP (Fig. 3.4C).

As expected, rat m-calpain displayed protease activity only when 2 mM Ca^{2+} is added and such activity was reduced when a cocktail of protease inhibitors were added (Fig. 3.4C; Boehmerle *et al.,* 2007). Rat m-calpain was included as a positive control and the corresponding data showed that a proteolytically active calpain could be purified using the optimized protocol. Additionally, the results also suggested that the non-activity

A


CAP5.5 FL
CAP5.5V FL
m-calpain
kDa
- + P S Ni2+
- + P S Ni2+
- + P S Ni2+
150
100
75
50
25

B


Suc-LLVY – N
H
S
N
N
S
COOH
Calpain I or II
Suc-LLVY + H2N
Aminoluciferin
Luciferase
ATP, Mg++, O2
Light
5105MA

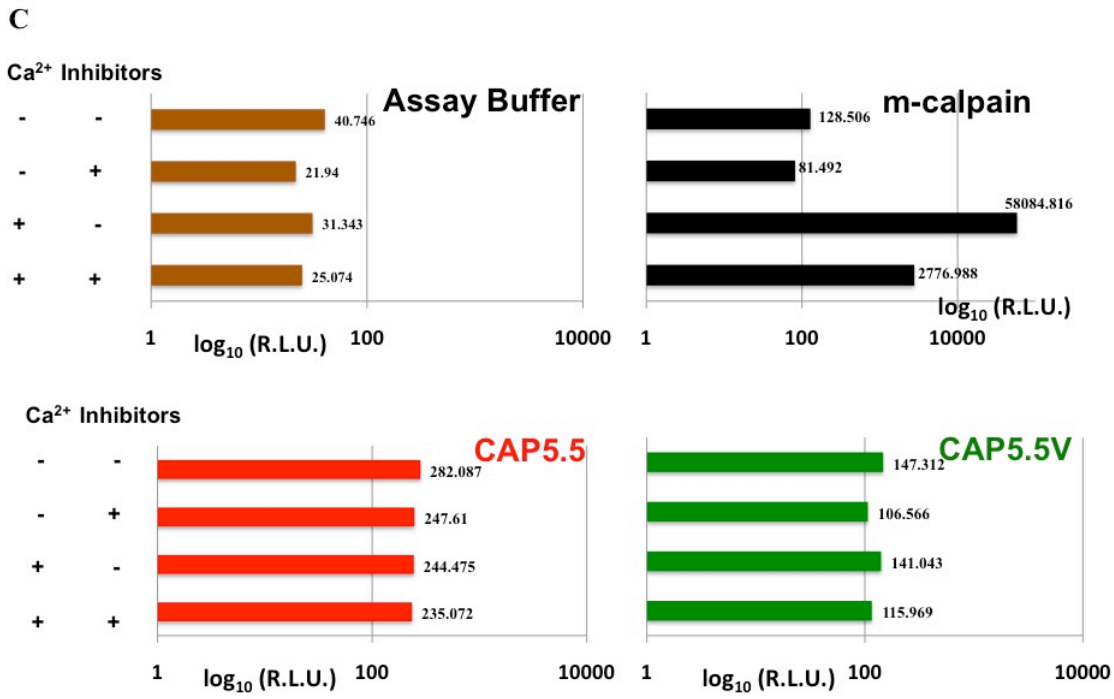


Figure 3.4. SDS-PAGE and protease assay profiles of CAP5.5 and CAP5.5V. A. Soluble CAP5.5-His_6 (120 kDa), CAP5.5V-His_6 (120 kDa) and rat m-calpain-His_6 (80 kDa) visualized on SDS-PAGE. Description of gel lanes: (-) Uninduced and (+) IPTG induced whole cell lysates; (P) insoluble pellet and (S) soluble supernatant from induced, lysed, centrifuged bacterial fractions. B. Schematic representation of the calpain activity assay, where the proluminescent substrate containing the Suc-LLVY is recognized by typical calpains. Following calpain cleavage, aminoluciferin (substrate for luciferase) is released, allowing for the luciferase reaction to occur, producing light. Image courtesy of Promega. C. Calpain activity assay of the assay buffer, m-calpain, CAP5.5 and CAP5.5V in the presence of 2 mM free Ca^{2+}, 0 Ca^{2+}, or with 2 mM free Ca^{2+} and a cocktail of protease inhibitors. R.L.U., relative luminescent units (arbitrary); F.L., full-length.

of CAP5.5 and CAP5.5V is not due to the presence of co-purifying contaminants that may inhibit calpain activity as all the recombinant calpains were expressed and purified in the same manner simultaneously. While the data may suggest a non-enzymatic, calcium independent role for CAP5.5 and CAP5.5V, which could be explained by the mutations in their catalytic triad and key calcium-binding residues, further biochemical characterization of the soluble proteins revealed an egregious heterodisperse aggregation (Section 3.3). Taken together, these results support an optimized method of obtaining soluble CAP5.5 and CAP5.5V, and corroborate the published expression and purification of an active m-calpain (Moldoveanu *et al.,* 2002).

3.4 Heterodisperse aggregation of CAP5.5 and CAP5.5V

Following nickel affinity chromatography, further chromatographic steps were performed in order to achieve pure and homogenous monomeric CAP5.5 and CAP5.5V required for biochemical and structural analyses. Using size exclusion chromatography on a Superdex 200 column, most of the protein (as determined by $A_{280\ nm}$ absorbance) eluted with the void volume (V_0). V_0 is the elution volume of molecules that are excluded from the gel filtration medium because they are larger than the largest pores in the matrix and pass straight through the packed bed. Based on the calculations from the molecular mass index standard, a *minor* peak corresponding roughly to the expected ~100 kDa molecular weight of monomeric CAP5.5V was also observed (Fig. 3.5). The same observations were made during the purification of CAP5.5.

The protein species of the V_0 and the predicted "monomeric" peaks were boiled in an SDS sample buffer containing either ß-mercaptoethanol (reducing) or no

BME (non-reducing). The effects of reducing agents such as BME were tested to assess the role of disulfide bridges in the elution of high-molecular species with the V_0 fraction. Reducing agents prevent intramolecular and intermolecular disulfide bonds from forming between cysteine residues of proteins. Upon analyzing the protein species of the V_0 and the predicted "monomeric" peaks on SDS-PAGE under reducing and non-reducing conditions, both the V_0 and the "monomeric" peak's gel lanes contained a band at ~120 kDa, corresponding to the expected size of CAP5.5V (Fig. 3.5 inset). Both CAP5.5 and CAP5.5V, despite having a predicted molecular weight of around 100 kDa, migrate slower on polyacrylamide gels by 20 kDa (Olego-Fernandez *et al*., 2009). Non-reducing SDS-PAGE of the V_0 fraction, "monomeric" fraction showed light smear band patterns above 250 kDa and collection in the lane wells (Fig. 3.5 inset arrow). Non-reducing SDS-PAGE of fractions eluted from a peak corresponding to 70 kDa did not exhibit such band patterns but also did not have a band at 120 kDa (CAP5.5V). This is highly suggestive of high molecular weight protein species that are unable to completely resolve and enter the gel. I also observed this phenomenon with the "monomeric" fractions suggesting either that CAP5.5V aggregated prior to SDS-PAGE analyses or that some of the "monomeric" fractions may be contaminated with high molecular weight species eluting from the void due to their peak proximity. Additionally, the 70 kDa fraction acts as a control to show that proteins can, in principle, enter the gel under non-reducing conditions without smearing or collecting in the wells.

However, these high molecular weight protein smears and collection in the lane wells disappeared when the same protein fractions were analyzed on a reducing SDS-PAGE, especially in the V_0 and "monomeric" fractions (Fig. 3.5 inset). Taken

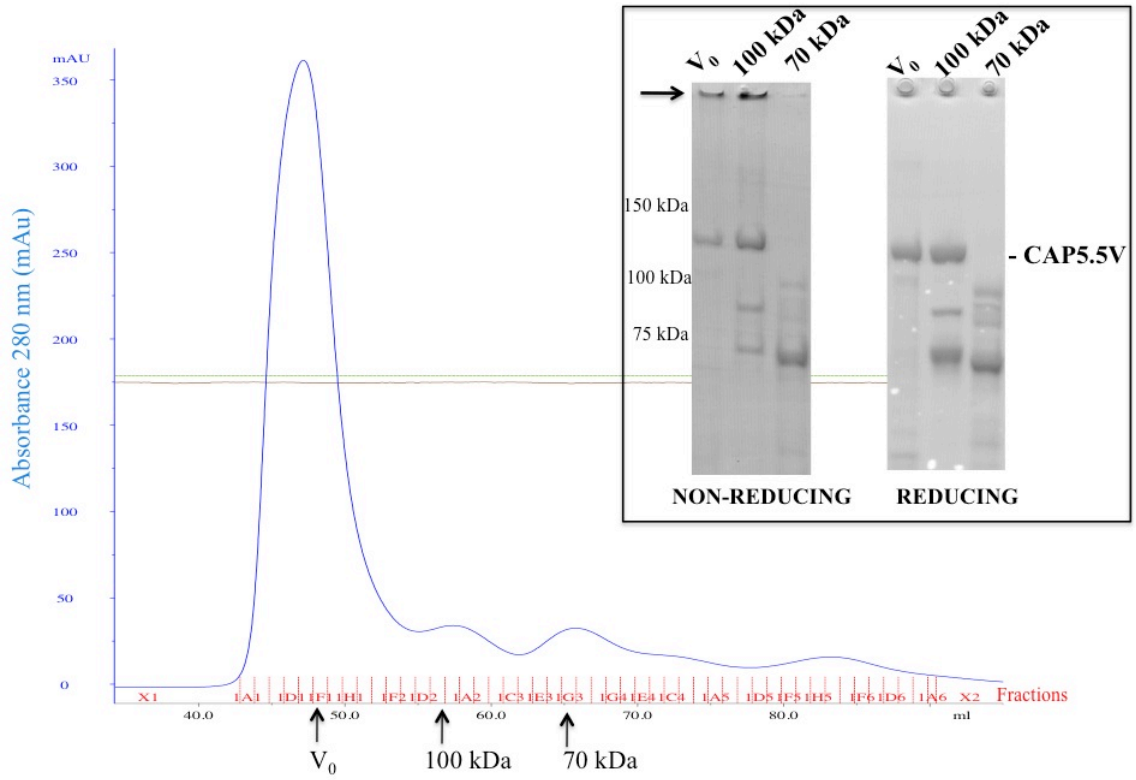


Figure 3.5 Aggregation of CAP5.5V. A representative 280 nm absorbance profile of size exclusion chromatography eluants using a Superdex 200 16/60 column. The void volume (V_0) and elution positions of water-soluble standard proteins are indicated below the profile. The inset shows SDS-PAGE analyses of proteins from V_0, 100 kDa and 70 kDa fractions under reducing (700 mM ß-mercaptoethanol) and non-reducing (no BME) conditions. The arrow indicates the collection of high-molecular weight species in the gel lane wells. Although, CAP5.5V elutes as 100 kDa, it runs as 120 kDa, as has been previously shown (Olego-Fernandez *et al.,* 2009).

together, these results are suggestive of a disulfide bridge dependent oligomerization pattern as reflected in the non-reducing SDS-PAGE. I hypothesized that under these non-reducing conditions, the incomplete reduction leaves some intra- or intermolecular disulfide bridges intact, making some SDS-binding domains unavailable to the detergent so that proteins are not saturated with SDS, thereby affecting their electrophoretic mobility. Additionally, CAP5.5 and CAP5.5V each contain ~20 cysteines and are hypothesized to *not* contain disulfide bridges between those cysteine residues. The cytosol, where the two are found, is highly reducing in its chemical nature, which strongly prohibits the oxidation of the sulfhydryl (-SH) groups, thereby preventing disulfide bonds between adjacent cysteine residues (Lopez-Mirabal, 2008). Therefore, it is likely that recombinant CAP5.5V forms disulfide bridges in solution that are not found under native cytosolic conditions, which may contribute to its instability and aggregation.

Typical calpains are generally heterodimers but mammalian calpains 3 and 13 have been found primarily in a homodimeric state (Ravulapalli *et al*., 2009). We used size exclusion chromatography coincident with in-line multi angle laser light scattering (MALLS) to determine the accurate molecular mass of the species that eluted with the V_0 under physiological conditions in order to determine if there was a specific order or pattern to the oligomerization such as dimerization etc (Fig. 3.6). MALLS yielded mass determinations of around 100 kDa as expected of the species in the "monomeric" CAP5.5V fraction but a heterodisperse mass determination ranging from 1000 kDa to 80, 000 kDa for species eluted in the V_0 (Figure 3.6). This suggested a disordered and heterodisperse aggregation of CAP5.5V molecules that may be driven by numerous factors including disulfide bridge formations.

3.5 Optimization of protein purification to resolve aggregation

Monodispersed proteins are those that are present in a single oligomeric state in solution, which are more likely to crystallize (D'Arcy, 1994; Zulauf & D'Arcy, 1992). In order to obtain monodisperse species of CAP5.5 and CAP5.5V, attempts were made to retroactively reverse aggregation by assaying this process on V_0 fractions from numerous runs of size exclusion chromatography. Analyses were done on V_0 fractions as most of the CAP5.5V eluted in these fractions, which were otherwise rather pure. Aggregation is inherently an equilibrium process and in the absence of autolysis, should be fully reversible (Pal *et al*., 2001). I hypothesized that as aggregation was a function of protein concentration, ionic strength and stabilizing protein-protein interactions, optimizing variables associated with these factors would help reverse aggregation of soluble CAP5.5 and CAP5.5V. Attempts to prevent aggregation *ante factum* were limited in scope as inclusion of high concentrations of any reducing agents (≥10 mM in my experience) in the lysis buffer reduces the bound nickel and eliminates the binding of hexahistidine tagged proteins during nickel affinity chromatography. Numerous rounds of expression and purification were done where the following variables were changed: buffer pH (5.1- 7.0); concentration (5, 10, 30 mM) and identity of reducing agents such as ß-mercaptoethanol (BME), dithiothreitol (DTT) and tris (2-carboxyethyl) phosphine (TCEP); detergents such as octyl glucopyranoside, triton X-100, triton X-114, IGEPAL (at concentrations below their respective critical micelle concentrations) and many others.

In one such representative experiment, CAP5.5V V_0 eluates (from size exclusion chromatography) were assayed for aggregation following treatment with

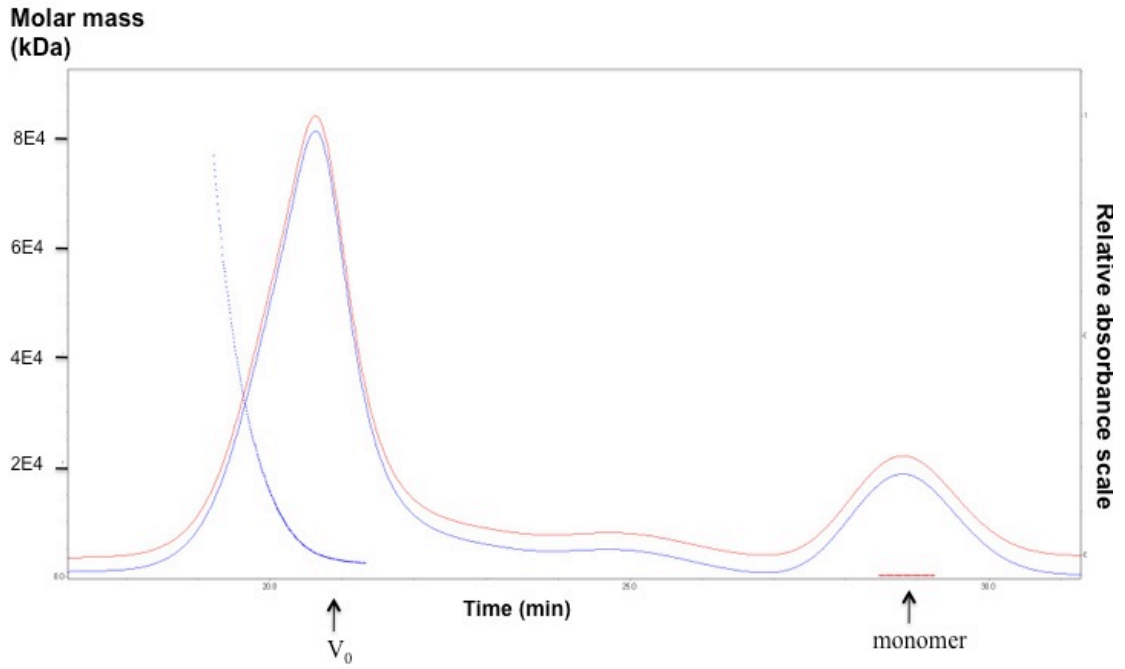


Figure 3.6. Molecular mass determination of CAP5.5V size exclusion chromatography eluates. Multi-angle laser light scattering of recombinant CAP5.5V is shown. The red and blue curves show the relative absorbance of two duplicates of the same CAP5.5V sample whereas the dotted lines show the determined mass. The MALLS data was generated by Steven Johnson.

different reducing agents of varying concentrations in buffers of differing pHs (Fig. 3.7). I assayed for aggregation by running the same treated protein samples on two types of protein gels: (1). An SDS-PAGE containing unboiled samples diluted in SDS sample buffer without any reducing agents. (2). A native PAGE containing unboiled samplesdiluted in buffer *sans* SDS and reducing agents even in the running buffer. The only difference between the two gels was the presence of SDS detergent in the SDS-PAGE gels. High molecular weight smearing and incomplete migration of proteins into the gel matrix were interpreted as proxies for aggregation.

Following incubation at 37ºC for 16 hours, SDS-PAGE analyses, irrespective of the conditions tested, did not show high molecular weight smears or proteins collecting in the lane wells as seen under native PAGE conditions (Fig. 3.7). In fact, a uniform band existed across all conditions corresponding to the molecular weight of monomeric CAP5.5V, which acted as a loading control for this experiment (Fig. 3.7A). A global comparison of the landscape of the two gels showed that proteins resolved and entered the gel matrix only under SDS-containing conditions. SDS is a negatively charged detergent used in electrophoresis to disrupt protein-lipid or protein-protein interactions (Neugebauer, 1990). Under SDS-containing conditions, the aggregation seen on native PAGE is lessened with the disruption of protein-protein interactions but the electrophoretic mobility of the protein sample is also undoubtedly aided by the negative charge of SDS.

In native PAGE, the electrophoretic mobility is dependent on both the protein's charge and its hydrodynamic size. Native PAGE analyses have been used to study processes that alter the conformation of proteins such as aggregation in numerous studies

where high molecular weight smearing and incomplete migration of proteins into the gel matrix were also interpreted as proxies for aggregation. (Arakawa *et al.,* 2001; Horan *et al*., 1995; Stegemann *et al.,* 2005). Although, I did not run a monodispersed protein in our experiments, I predicted, based upon on those studies, that the proxies used for aggregation would not be present in a non-aggregated protein sample. Native PAGE analyses of CAP5.5V treated under the various conditions all showed a resolved protein band, which I take to be monomeric CAP5.5V; high molecular weight smears, and protein collection in the lane wells (Fig. 3.7B). This observation was made of all conditions albeit with slight differences. For instance, one might notice that at pH 5.0 as compared to pH 7.0, the high molecular weight smearing is more pronounced and the resolved band less intense.

Based on these parameters, I hypothesized that at a pH closer to the predicted pI of CAP5.5V (4.2) such as at pH 5.0, the protein carries less charge than at a more alkaline pH 7.0 to migrate as well into the gel. Therefore under acidic conditions, less protein may be entering the gel matrix thereby explaining the less prominent high molecular weight smears. Nevertheless, in the native system, the indicators for aggregation remained regardless of the variation in pH, reducing agents and their concentrations even though the native PAGE gel was run for a longer period of time. Taken together, the data revealed the complexity of CAP5.5V aggregation, which I posit is not only mediated by disulfide bridges formation but also some other unknown variable(s).

In another representative experiment, CAP5.5V was expressed and purified as before except the nickel affinity chromatography eluates were collected and incubated in

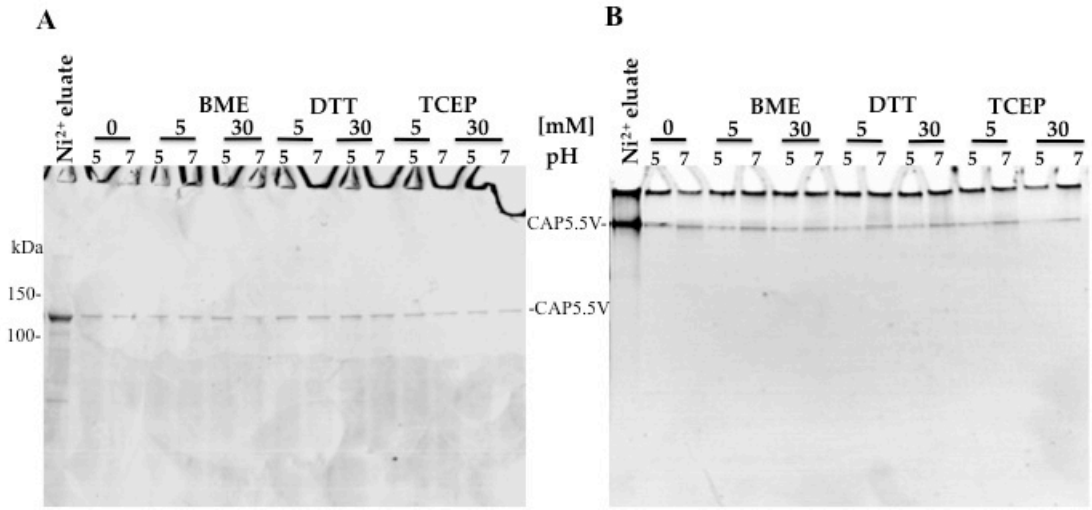


Figure 3.7. Assaying for CAP5.5V aggregation under varying pH and reducing conditions. A. 10% SDS-PAGE of unboiled CAP5.5V nickel chromatography eluate as well as the void volume eluate (V_0) in a sample buffer (62.5 mM Tris-HCl, pH 6.8; 40% glycerol; 0.01% bromophenol blue; 2% SDS) under various conditions. The running buffer contained SDS. **B.** 10% native PAGE (non-boiled; sample buffer consisting of 62.5 mM Tris-HCl, pH 6.8; 40% glycerol; 0.01% bromophenol blue) of the same protein samples as in A. The running buffer did not contain any SDS. The protein samples had been incubated for 16 hours at 37ºC. ß-mercaptoethanol (BME), dithiothreitol (DTT) and tris (2-carboxyethyl) phosphine (TCEP).

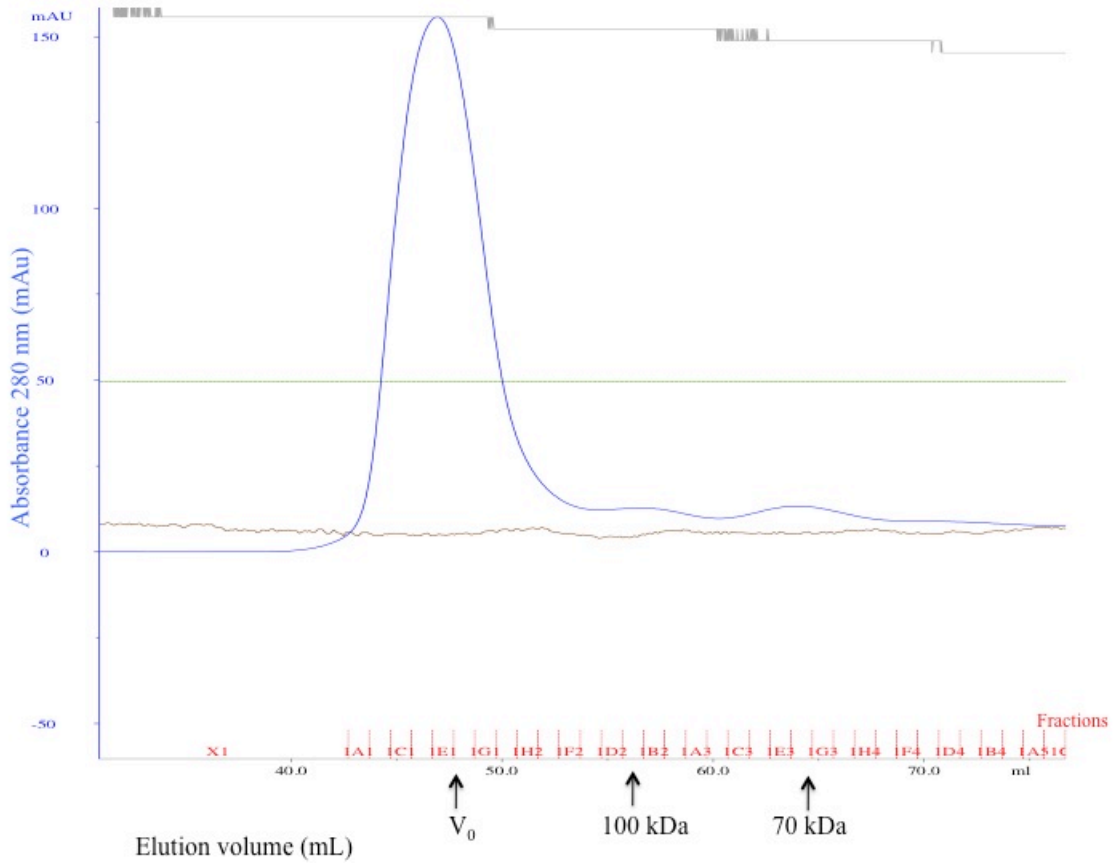


Figure 3.8. Size exclusion chromatography elution profile of recombinant CAP5.5V following iodoacetamide treatment. The 280 nm absorbance profile of eluants using a Superdex 200 16/60 column is shown. The void volume (V_0) and elution positions of water soluble standard proteins are indicated below the profile.

buffer containing iodoacetamide, an irreversible acetylating agent that replaces the thiol group of free cysteines with an ethanamide group, thereby inhibiting any disulfide bridge formation that should not be present in a cytosolic protein like CAP5.5V (Smythe, 1936). Upon proceeding to size exclusion chromatography, however, CAP5.5V still eluted with V_0 and no major shift from the V_0 towards the "monomeric" peak was noticeable (Fig. 3.8). Assuming that iodoacetamide-based acetylation of cysteines were successful –I have not confirmed this by mass spectrometry- these results imply a multifactorial cause of CAP5.5 and CAP5.5V soluble aggregation, as the aggregation persisted despite the chemical inhibition of disulfide bond mediated complex formation.

Aggregation of calpains has been frequently noted and impeded structural characterization of typical calpains for many years (Yoshizawa *et al.*, 1995; Elce *et al.*, 1997). In the case of the dimeric rat m-calpain, Ca^{2+} activation led to subunit disassociation, exposing the hydrophobic dimerization interface causing oligomerization (Pal *et al.*, 2001). The structures of mammalian calpains were solved using ethylenediaminetetraacetic acid (EDTA) to chelate calcium or endogenous inhibitors (Moldoveanu *et al.,* 2002; Hanna *et al.*, 2008). The purification protocol described here also used EDTA to chelate free calcium ions even though CAP5.5 and CAP5.5V lack three out of the five conserved acidic calcium-binding residues in the catalytic domain of calpains. Unable to resolve CAP5.5 and CAP5.5V protein aggregation, I proceeded to further characterize CAP5.5V using a series of truncation constructs.

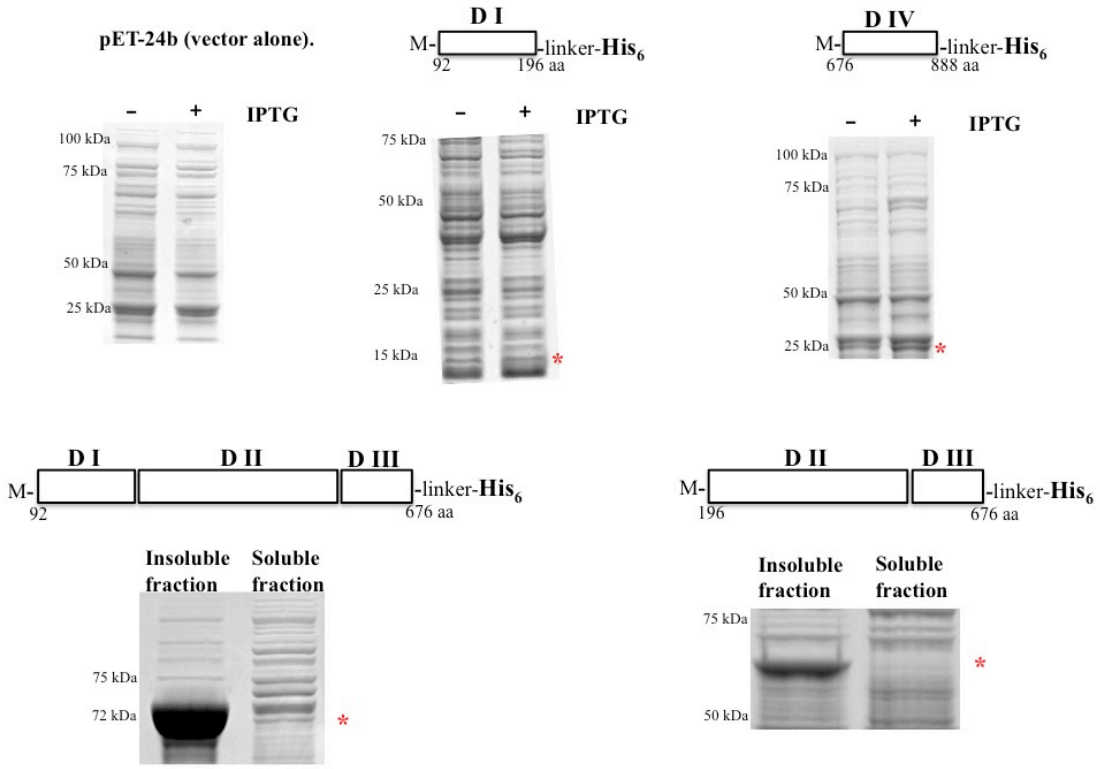


Figure 3.8. SDS-PAGE analyses of protein expression. SDS-PAGE analyses of bacterial cell fractions before (-) and after (+) 0.5 mM IPTG is shown for pET-24b (the parental vector), constructs expressing CAP5.5V domain I (14 kDa) and domain IV (24 kDa). For constructs expressing domains I-III (71 kDa) and domains II-III (67 kDa), the insoluble and soluble fractions following the lysis of induced bacterial cells are shown. The expected molecular weight is denoted by a red asterisk (*). Linker, VDKLAAALE.

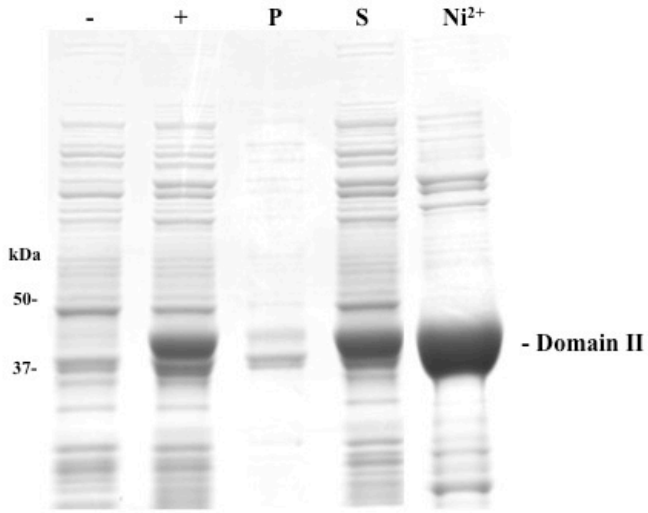


Figure 3.9. Soluble expression of CAP5.5V domain II visualized by SDS-PAGE. Description of gel lanes: (-) Uninduced and (+) IPTG induced bacterial cell lysates; (P) insoluble pellet and (S) soluble supernatant from induced, lysed, centrifuged bacterial fractions. The predicted molecular weight for CAP5.5V domain II is 40 kDa.

3.6 Expression and purification of CAP5.5V domain II

A majority of the CAP5.5V truncation constructs either failed to produce discernible levels of protein expression (DI, DIV) or fractionated with the insoluble pellet (DI-III, DII-III) (Fig. 3.8). However, one construct encoding CAP5.5V domain II (196-538), which is homologous to the protease core domain of typical calpains, expressed upon IPTG induction and was soluble (Fig. 3.9). Purification steps included a round of nickel affinity chromatography followed by one of size exclusion chromatography. Size exclusion chromatography revealed two prominent peaks: the taller one corresponding to the void volume and a smaller peak that was predicted to contain ~40 kDa molecular weight proteins, as determined by the elution positions of water-soluble standard proteins (Fig. 3.10). CAP5.5V domain II is predicted to have a molecular weight of 40 kDa. Most of CAP5.5V domain II eluted with the V_0 as judged by SDS-PAGE analyses of the V_0 and "monomeric" fractions (Fig. 3.10 inset). We further analyzed monomeric CAP5.5V domain II under physiological conditions by multi-angle laser light scattering (MALLS) and found that most of the protein still eluted as a peak with a molecular mass of around 40 kDa corresponding to monomeric domain II. However, we found a small peak with molecular mass ranging from 250 to 350 kDa (Fig. 3.11 blue line). These results not only suggested that domain II aggregated just as was seen with the full-length versions of CAP5.5 and CAP5.5V but also that purified monomeric species tended to aggregate over time in solution.

Building upon prior optimization trials of CAP5.5 and CAP5.5V aggregation, a series of new variables were tested including one that varied the ionic strengths of buffers. I hypothesized, based upon the homology models discussed in

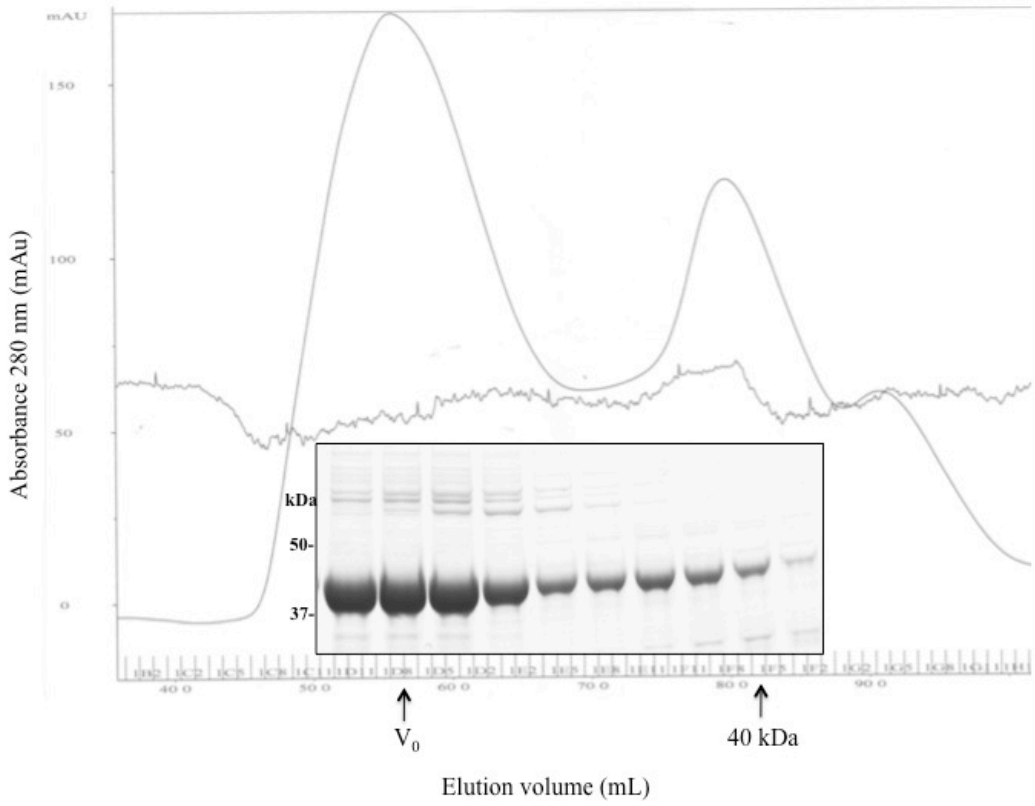


Figure 3.10. Elution profile and corresponding SDS-PAGE analyses of the fractions. Size exclusion chromatography was done on an S200 16/60 column and the absorbance 280 nm profile of CAP5.5V domain II (40 kDa) elution is shown. An SDS-PAGE gel (inset) shows the protein composition roughly corresponding to the location of the gel lane on the chromatograph.

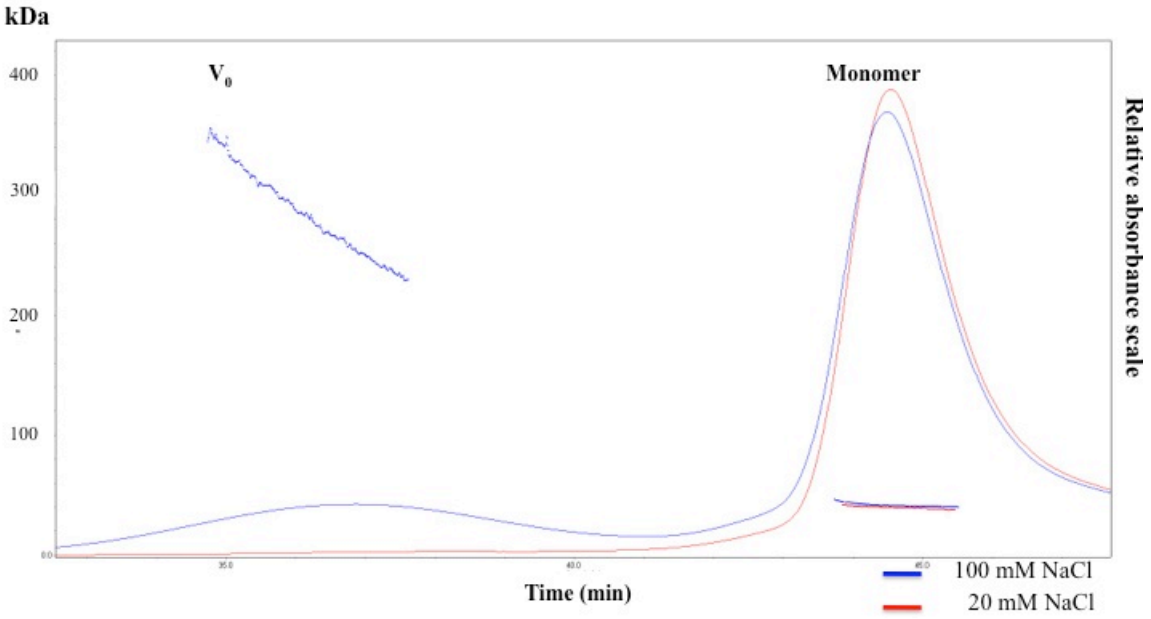


Figure 3.11. Molecular mass determination of CAP5.5V domain II under two conditions of varying ionic strength. Multi-angle laser light scattering of recombinant CAP5.5V domain II (40 kDa) is shown. The red curve shows the absorbance profile of domain II in buffer containing 20 mM NaCl and the blue curve show that for domain II in buffer containing 100 mM NaCl. The dotted lines show the determined mass. The MALLS data was generated by Steven Johnson.

Section 3.2, that the interaction interface between domain II and domain III may be hydrophobic. Therefore, when domain II is expressed alone, to reduce that hydrophobic patch from interacting with other such patches on the protein, decreasing the ionic strength would promote an increase in the water solvation layers around a protein molecule, thereby mitigating aggregation. When we compared the MALLS elution profiles of aggregated CAP5.5V domain II "monomeric" fractions that had been dialysed into either a high (100 mM NaCl) or low salt (20 mM NaCl) buffer overnight at 4ºC, we observed, under low salt conditions, a reduction in the peak corresponding to aggregated material and a slight increase in the peak corresponding to monomeric species (Fig 3.11). This suggested that buffers with low ionic strength might retroactively inhibit hydrophobic interaction-mediated aggregation of CAP5.5V domain II. This new protocol was also used for the purification of full-length CAP5.5V but to no avail suggesting a far more complex and multi-factorial cause of aggregation for the full-length protein.

With such modifications namely the use of 20 mM salt in all buffers (the fully optimized protocol is in Chapter 2), CAP5.5V domain II was expressed and purified to near homogeneity as determined by Sypro-Ruby staining of protein fractions (Fig. 3.12A). This material was then used for biochemical and structural analyses.

3.7 Crystallization trials of CAP5.5V domain II

Crystallization screens were set up robotically and various commercial crystallization screen kits were used. A few hundred sets of screening conditions, varying in buffer systems and precipitants, were tested for sitting drop vapor diffusion crystallization of recombinant CAP5.5V domain II by sparse matrix sampling. Under the

current conditions, no crystal growth was observed and therefore, will require further optimization in the future.

3.8 CAP5.5V domain II has no detectable protease activity

CAP5.5V has been shown to be essential in the cytoskeletal remodelling of the bloodstream forms of *T. brucei* (Olego-Fernandez *et al*., 2009). Whether, such a role is dependent upon an intrinsic proteolytic mechanism remains of great interest due to its classification in the calpain superfamily, which is a calcium-regulated cysteine protease (Ersfeld *et al.,* 2005). Additionally, elucidation of its function may unravel new avenues of drug development against this essential protein for African sleeping sickness.

The protease core domain of rat CAPN1 was shown to have weak but detectable proteolytic activity upon calcium loading. This was observed even in the absence of the other domains, suggesting that the protease core acts co-operatively upon calcium activation to form an independent and functional protease unit (Moldoveanu *et al.,* 2002; Moldoveanu *et al.,* 2004a). Comparison of the proteolytic activity of the protease core and the whole calpain revealed that while both enzymes had similar initial reaction rates, the protease core maintained its enzymatic activity for a longer duration (Campbell & Davies, 2012). I was interested in whether CAP5.5V domain II, which corresponds to the catalytic domain of typical calpains, exhibited similar activity.

The active sites residues of kinetoplastid calpains were predicted following sequence alignments and in some cases, homology modelling with typical calpains (Ersfeld *et al.,* 2005). A majority of these calpains including CAP5.5 and CAP5.5V possess catalytic triad residues that deviate from the canonical catalytic triad residues of

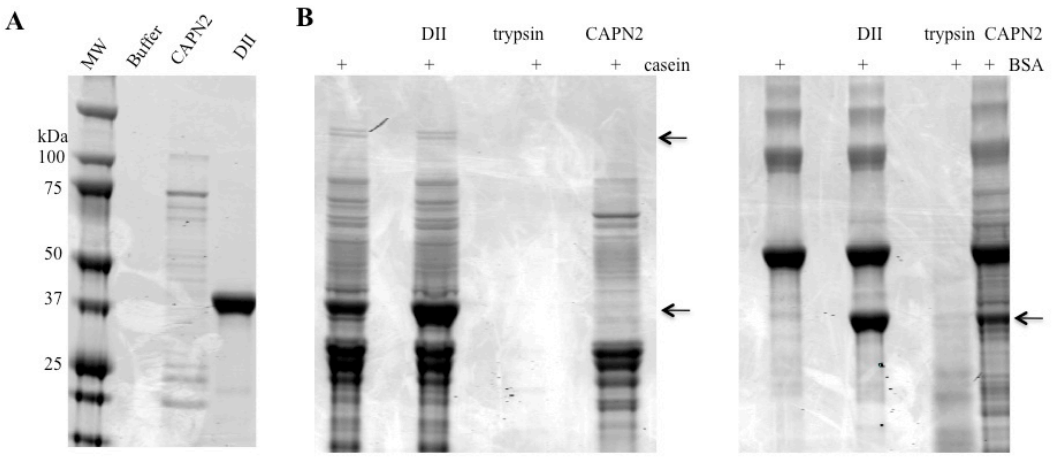


Figure 3.12. CAP5.5V domain II has no protease activity. A. The purity of recombinant CAP5.5V domain (40 kDa) and rat CAPN2 (80kDa) was assessed. **B**. Activity of domain II was assayed by SDS-PAGE analyses of fragments following incubation of recombinant domain II with commercial-grade substrates such as casein and bovine serum albumin (BSA). As positive controls, trypsin and the nickel eluate of recombinant rat m-calpain (CAPN2) was used. Equal concentration (10 µg) of substrates and recombinant calpains were used. The expected molecular weights of BSA and casein are 66 kDa and 24 kDa respectively, however, various isoforms exist which might explain the hetereogenity on the gels.

cysteine proteases (C-H-N), with CAP5.5/CAP5.5V possessing a potential serine nucleophile and tyrosine instead of the basic histidine (S-Y-N). Even though CAP5.5V has major substitutions in its catalytic triad residues, not all atypical active site configurations preclude catalytic activity. For example, the poliovirus-type picornain C3 peptidases are classified as serine proteases but operate with a cysteine nucleophile in cleaving viral proteins (Sarkany & Polgar, 2003). However, no cysteine proteases operating with a serine nucleophile have been hitherto observed with proteolytic activity.

CAP5.5V domain II's proteolytic activity was tested in an assay in which equal concentrations (10 μg) of commercial-grade casein or bovine serum albumin (BSA) was incubated with CAP5.5V domain II or rat CAPN2 or trypsin for 60 minutes at 37ºC. Proteolytic degradation of these universal protease substrates were monitored by SDS-PAGE analyses as was done in the biochemical characterization of other calpains (Wang *et al.,* 2003; Hata *et al.,* 2001).

Purity of recombinant CAP5.5V domain II and rat CAPN2 was assessed by visualization of 10 μg of these proteins along with the calcium-containing incubation buffer on SDS-PAGE (Fig. 3.12A). The buffer appears to have no proteinaceous contaminants while CAP5.5V domain II is quite pure as judged by the lack of any other visible bands. Additionally, one notices little autoproteolytic activity for CAP5.5V domain II (Fig. 3.12A). Full-length rat CAPN2 was partially purified as judged by its SDS-PAGE profile. It has already been shown to possess proteolytic activity that is not due to a co-purifying contaminant (Fig. 3.4), and was used in this experiment as a positive control.

Proteolytic degradation by trypsin of both substrates was reflected in the light smear and that of rat CAPN2 by the disappearance of high molecular weight proteins (Fig. 3.12B arrows). However, upon incubating casein or BSA with CAP5.5V domain II, SDS-PAGE profiles revealed no differences in the band patterns, thereby suggesting little degradation of these substrates (Fig. 3.12B). Given the high concentration (10 µg) of both substrates and recombinant calpains used, I believe the presence of any residual protease activity of CAP5.5V domain II would be detected in this analysis. While, it cannot determined if the purified domain has properly folded, its solubility and monodispersity in solution is high suggestive of it being so. I, therefore, suggest that CAP5.5V domain II is not proteolytically active as determined under the present conditions. By extension, this points to a probable non-enzymatic role for CAP5.5 and CAP5.5V in *T. brucei,* possibly in regulatory processes based on observations made for typical, active calpains but also for atypical, non-active enzymes such as kinases and proteases (Todd *et al.,* 2002; Pils & Schultz, 2004; Ersfeld *et al.,* 2005). However, detailed functional studies will be necessary to sufficiently address these issues.

3.9 Summary

In this study, the structural domain boundaries of CAP5.5 and CAP5.5V were delineated in order to create a series of expression constructs for biochemical and crystallographic analyses. Aggregation of soluble full-length CAP5.5 and CAP5.5V and the inability to reverse this process despite further optimization led us to pursue the expression of CAP5.5V's individual domains or combinations thereof. Following the determination and reversal of disulfide bond and hydrophobic-mediated aggregation,

monodispersed and purified CAP5.5V domain II was used for physicochemical characterization. CAP5.5V domain II was shown to have no detectable protease activity under the present conditions, suggesting a non-enzymatic and possibly, regulatory role in *T. brucei.*

4. Characterization of CAP5.5 and CAP5.5V in *T. brucei*

The work encompassed in figures 4.8 and 4.9B- 4.9D was contributed by Samuel Dean (University of Oxford).

4.1 Introduction

RNAi-based functional characterization of CAP5.5 and CAP5.5V has revealed their crucial roles for the correct cell morphogenesis of procyclic and bloodstream forms of *T. brucei* respectively (Olego-Fernandez *et al.,* 2009). These differentially expressed calpains are associated with the cytoskeleton and CAP5.5 has been shown to be dually acylated with myristic acid and palmitic acid (Hertz-Fowler *et al.,* 2001; Olego-Fernandez *et al.,* 2009). For a fuller introduction to CAP5.5 and CAP5.5V, please refer to Chapter 1.

Cellular localization analyses of mono- and diacylated proteins suggest that while dual acylation in itself appears to elicit flagellar membrane association in kinetoplastid parasites, this signal can be superseded by additional cues (Emmer *et al.,* 2009). For example, CAP5.5 and CAP5.5V localize to the whole cell body but are excluded from the flagellum (Hertz-Fowler *et al.,* 2001; Olego-Fernandez *et al.,* 2009). In this case, the "cue" is hypothesized to be mediated by their biochemical association with the sub-pellicular cytoskeleton (Hertz-Fowler *et al.,* 2001).

My aims in this chapter were two-fold: firstly, to corroborate and further characterize the function of CAP5.5 by over-expressing the full-length protein at two different levels for dominant-negative analyses. Secondly, to investigate the molecular determinants in CAP5.5's sequence such as motifs or domains, which are involved in targeting the dually acylated calpain to the sub-pellicular cytoskeleton. This was achieved

by the ectopic expression and subsequent characterization of a series of CAP5.5 deletion mutants.

4.2 Expression constructs

Based on the delineation of CAP5.5 domain boundaries outlined in Chapter 3, the series of constructs used in this chapter were designed (Fig. 4.1). The ectopic expression vectors used in this chapter were either pDex577 and/or pDex777 (Kelly *et al.,* 2007; Poon *et al.,* 2012; for vector maps, see chapter 2). The primary difference between the two vectors is the 3' intergenic regions of the induced gene; pDex577 has an aldolase 3' UTR and pDex777 has paraflagellar rod 2 (PFR-2) 3' UTR immediately downstream of the GFP-coding sequences.

Given the polycistronic organization of the trypanosomatid genome, most of the regulation is largely post-transcriptional where mRNA levels are determined by modulating mRNA stability and processing. Such regulation is mediated by *trans*-acting factors that recognize *cis*-acting sequences in the 3' UTR of mRNAs (reviewed in Clayton, 2002; Ouellette & Papadopoulou, 2009). As such, owing to the identity of the 3'UTR, transgene expression is much greater when using pDex777 as compared to pDex577 (Poon *et al.,* 2012).

In addition to the full length CAP5.5 (1-853), the first 6 (N6), 37 amino acids (N37), domain I-III (1-659), domain I-II (1-521) and domain I (1-190) were engineered into pDex577 vectors fused to a C-terminal YFP::Ty tag for over-expression in procyclic cells. The full-length CAP5.5 fused to a C-terminal YFP::Ty (pDex777) was also

expressed in procyclic parasites to allow a comparison in phenotype when ectopically expressed at a significantly higher level.

The aforementioned constructs also expressed a GS_{10} linker between the protein of interest and GFP/YFP::Ty (Fig. 4.1). To achieve this, I subcloned a DNA fragment encoding that linker region into pDex vectors with the hypothesis that flexible linkers facilitate the independent folding of the two proteins (Chen *et al.,* 2013). Additionally, owing to the limited choices of restriction enzyme sites (HindIII and SpeI only) for subcloning DNA fragments into pDex vectors to engineer C-terminal GFP tagged proteins and the finding of several native HindIII sites in CAP5.5/CAP5.5V, an extended multiple cloning region was also subcloned in frame to the HindIII and SpeI sites of pDex577 and pDex777. For the sequence of this fragment, please refer to Table 2.1.

The full-length CAP5.5 (1-853) and full length CAP5.5V (1-888) were over-expressed using pDex777 vectors with an N-terminal GFP tag in both procyclic and bloodstream forms (Fig. 4.1). Ty::YFP::Ty alone was used as a control and expressed in both vectors. For the rationale behind the design of each construct, please refer to the corresponding sections. Expression of fusion proteins in the transfectants was determined by native fluorescence and/or indirect immunofluorescence analysis, following biochemical corroboration by western blots.

4.3 Over-expression of CAP5.5 in procyclic *T. brucei*

To analyse the function of CAP5.5 in procyclic forms of *T. brucei*, a cell line over-expressing the full-length CAP5.5 with a C-terminal YFP::Ty tag was generated

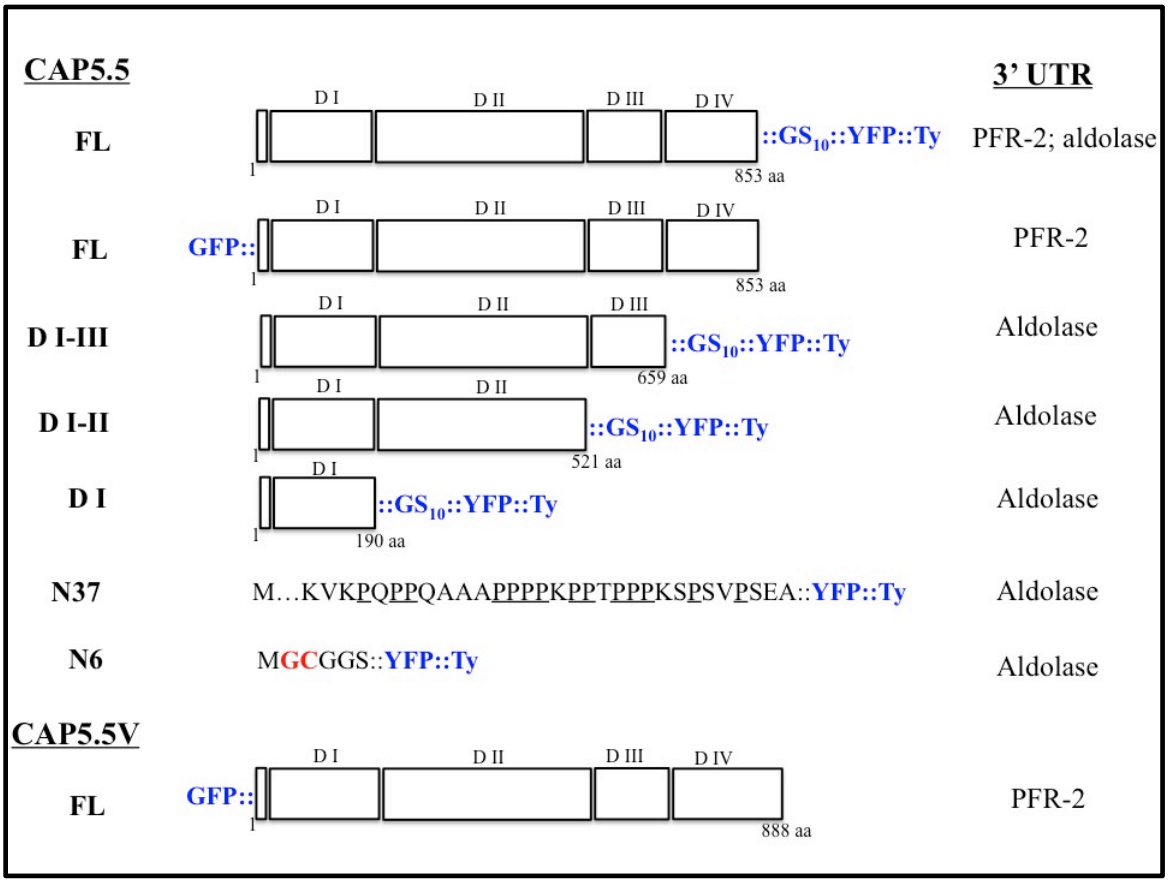


Figure 4.1. Schematic representation of the expression constructs used in this chapter. Two over-expression vectors in *T. brucei* were used in this study: pDex577 which uses the aldolase 3'UTR and pDex777 which utilizes the paraflagellar rod 2 (PFR-2) 3'UTR (Poon *et al.,* 2012). The two vectors differ only in their intergenic regions. The tags (GFP/YFP and Ty epitope) are highlighted in blue. GS_{10} is a linker region comprising of 5 glycines (G) and 5 serines (S; sequence in Table 2.1). N6 and N37 refer to constructs over-expressing the first 6 or 37 amino acids of CAP5.5, where the amino acids in red are the predicted myristoylation/palmitoylation residues. The underlined residues highlight the proline-rich region in that short N-terminal stretch of CAP5.5.

using the pDex577 vector (CAP5.5::YFP::Ty). Additionally, as CAP5.5's cellular localization has already been determined, full-length CAP5.5 over-expression construct was engineered with a GS_{10} linker::YFP::Ty exactly as the series of deletion mutants to serve as a control to ensure that the ectopic CAP5.5 recapitulates the characteristics of wild-type CAP5.5 (Hertz-Fowler *et al.,* 2001). We were also interested in determining how CAP5.5 gets built into the sub-pellicular cytoskeleton and whether such incorporation is cell cycle regulated. Inducible over-expression of CAP5.5 was confirmed by western blot analyses using antibodies against CAP5.5, which showed a marked but comparatively lower CAP5.5::YFP::Ty expression than native CAP5.5 after 24 hours of tetracycline induction (Fig.4.2A). Over-expression of CAP5.5 did not lead to a visible effect on the growth rate of the population over five days when compared with the non-induced population (Fig. 4.2B). This suggests that the procyclic-specific CAP5.5, at these expression levels, does not lead to a noticeable dominant negative phenotype when over-expressed in procyclic forms.

Localization analyses by native fluorescence microscopy revealed that YFP::Ty tagged CAP5.5 localized, exactly as the wild-type version, to the cell body and was excluded from the flagellum after 24 hours of induction (Fig. 4.2C; Hertz-Fowler *et al.,* 2001). One can also notice an accumulation around the periphery of the cell, which is compatible with the model of this diacylated calpain being associated with the cell body membrane. CAP5.5/CAP5.5V are associated with the cytoskeleton and are therefore resistant to non-ionic detergent extraction *in vitro* and *in situ* (Sherwin & Gull, 1989a). Upon detergent extraction, CAP5.5::YFP::Ty expression remained, suggesting it is

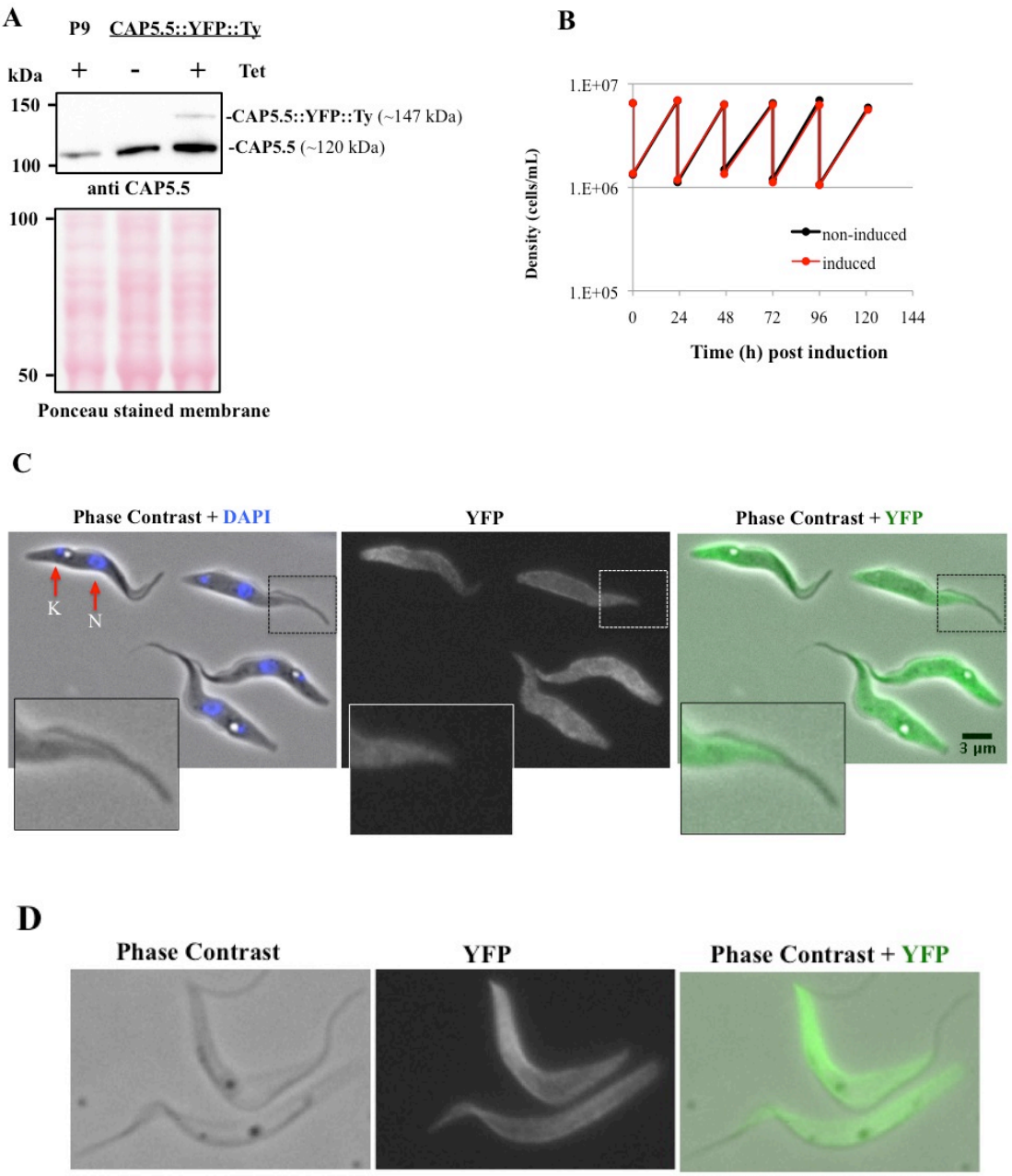


Figure 4.2. Over-expression of CAP5.5::YFP::Ty in procyclic cells. A. Western blot of 3.5 x10^6 cells/lane of parental (P9), cells over-expressing CAP5.5::YFP::Ty (pDex577 vector) under tetracycline induced (+) or non-induced (-) conditions using anti-CAP5.5 antibody. Ponceau-stained membrane is for loading control. **B.** Growth curve of induced and non-induced cells. **C.** Phase contrast and fluorescence micrographs of 2% PFA-fixed of cells over-expressing CAP5.5::YFP::Ty after 24 hours of induction. K and N (red arrow) refer to the kinetoplast and nucleus. DNA was labelled with DAPI (cyan). **D.** Cells from the same induction as C were treated with detergent (cytoskeleton), fixed in methanol and their localizations assessed by native fluorescence microscopy. Scale bar, 3 µm.

associated with the sub-pellicular cytoskeleton in the same way as native CAP5.5 (Fig. 4.2D; Hertz-Fowler *et al.,* 2001).

4.4 Temporal expression and localization of CAP5.5

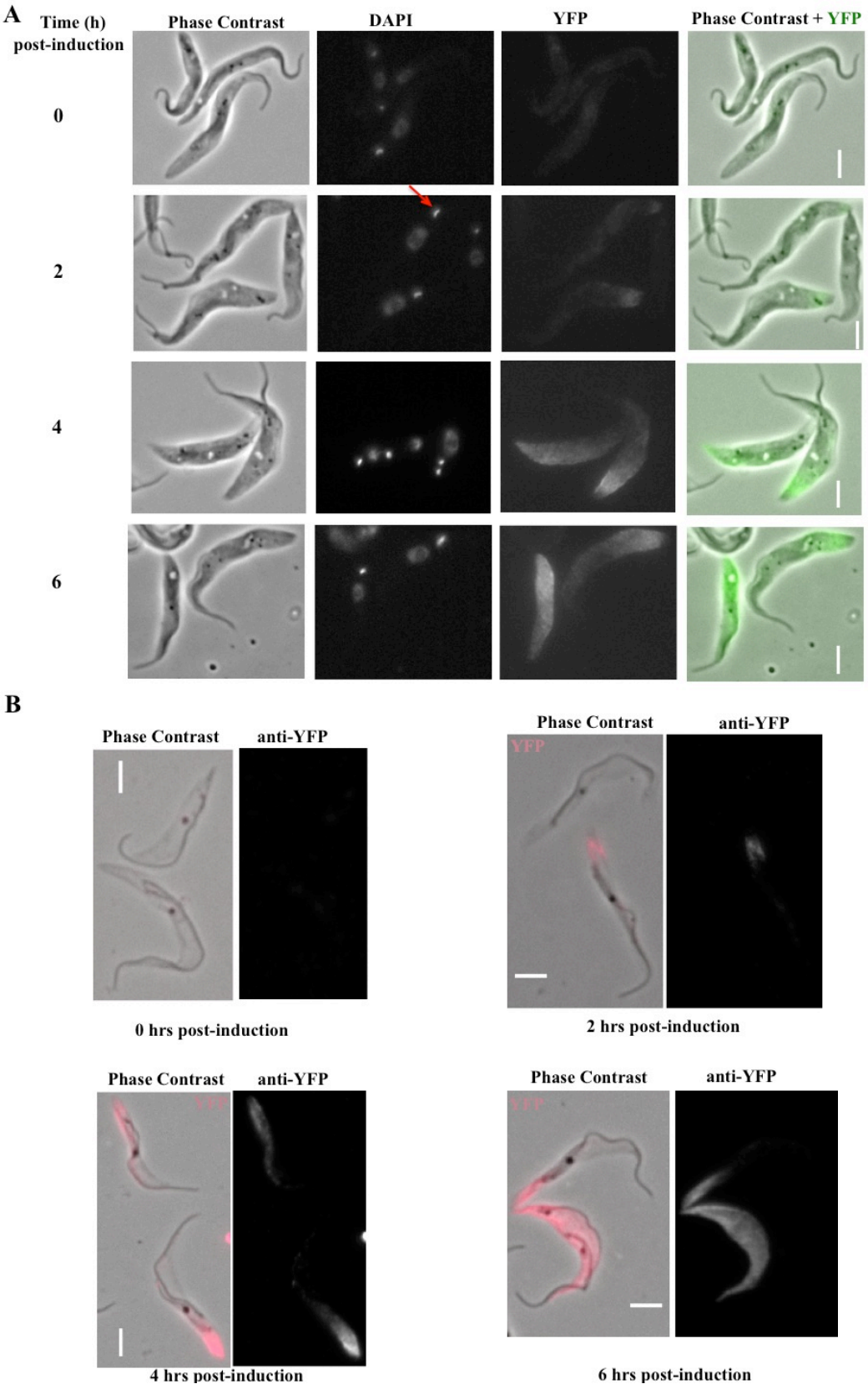
Given the lack of a dominant negative effect following CAP5.5 over-expression, whether it be the native CAP5.5-like localization of CAP5.5::YFP::Ty or the lack of a growth defect, the data suggested that the fusion protein is a good model for its wild-type counterpart. Therefore, we further analyzed cells over-expressing CAP5.5::YFP::Ty in order to better understand CAP5.5 expression, how CAP5.5 gets built into the sub-pellicular cytoskeleton and whether such an event is cell cycle regulated.

In order to monitor the temporal localization of CAP5.5::YFP::Ty upon tetracycline induction, a time-course analysis of cells at 0, 2, 4, and 6 hours following induction was done. Fluorescence micrographs revealed that at 2 hours post-induction, the YFP signal was restricted to the posterior end of the cell, with the signal becoming more widespread gradually across the cell body by 4 and 6 hours post-induction (Fig. 4.3A). Detergent-extraction (cytoskeletons) of induced cells at those same time points by indirect immunofluorescence microscopy revealed a corresponding localization to the whole cells, with CAP5.5::YFP::Ty associated with the posterior ends of the sub-pellicular cytoskeleton at 2 hours post-induction and later (Fig. 4.3B). In order to ascertain the posterior ends of the induced cells at 2 hours post-induction had actively extending microtubules, a YL1/2 monoclonal antibody, that recognizes tyrosinated α-tubulin was used (Kilmartin *et al.,* 1982). Labelling of this antibody, which stains dynamic microtubules in the trypanosomal cytoskeleton, was observed in the YFP-

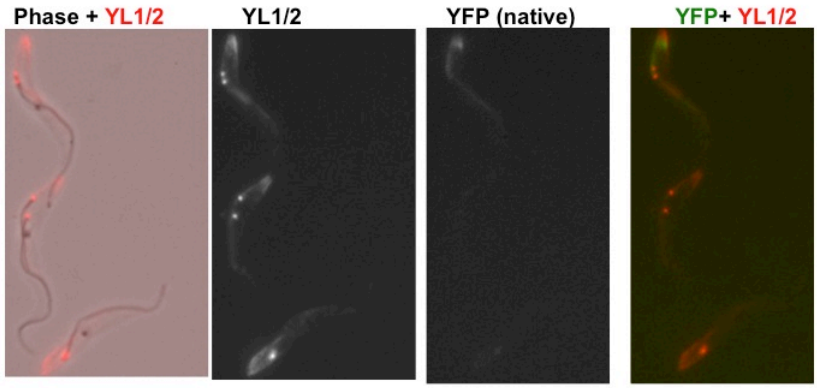
positive cells at 2 hours post-induction, albeit posterior to the CAP5.5::YFP::Ty signal (Fig. 4.3C). Although, there is an overlap in the localization of CAP5.5::YFP::Ty and YL1/2 antibody staining, they do not perfectly co-localize. This may mean that the incorporation of CAP5.5::YFP::Ty is not perfectly co-ordinated with the incorporation of new microtubules.

The microtubule-based cytoskeleton of trypanosomes is oriented such that the fast-growing, dynamic and plus ends of the microtubules are located at the posterior end, where initiation of cell elongation takes place during the growth phase of the cell cycle (Sherwin, & Gull, 1989b; Robinson *et al.,* 1995). This observation complements the RNAi phenotype of CAP5.5, where loss of CAP5.5 occurs first from the posterior cell end to eventually result in aberrant cytokinesis (Olego-Fernandez *et al.,* 2009). Taken together, this implies that CAP5.5 and presumably, other microtubule-associated proteins (MAPs) are added to the newly re-modelled sub-pellicular microtubule array during cell elongation at its dynamic posterior ends. Some MAPs like CAP5.5, then move from the posterior to the anterior ends of the cell, having been built in from the cell's posterior ends to eventually associate with the entire sub-pellicular cytoskeleton as is observed in the later stages of the CAP5.5::YFP::Ty over-expression.

To assess if CAP5.5::YFP::Ty incorporation into the sub-pellicular cytoskeleton at these early time points following induction is cell cycle regulated, I determined the approximate positions of individual cells in the cell cycle by 4, 6-diamidino-2-phenylindole (DAPI) staining of cells at each time point (0, 2, 4, 6 hours post induction). The temporal order of the key events in the procyclic cell division cycle is well characterized and readily determined by DAPI staining of the mitochondrial (kinetoplast)


A
Time (h) post-induction
Phase Contrast
DAPI
YFP
Phase Contrast + YFP
0
2
4
6
B
Phase Contrast
anti-YFP
YFP
0 hrs post-induction
2 hrs post-induction
4 hrs post-induction
6 hrs post-induction

C



D

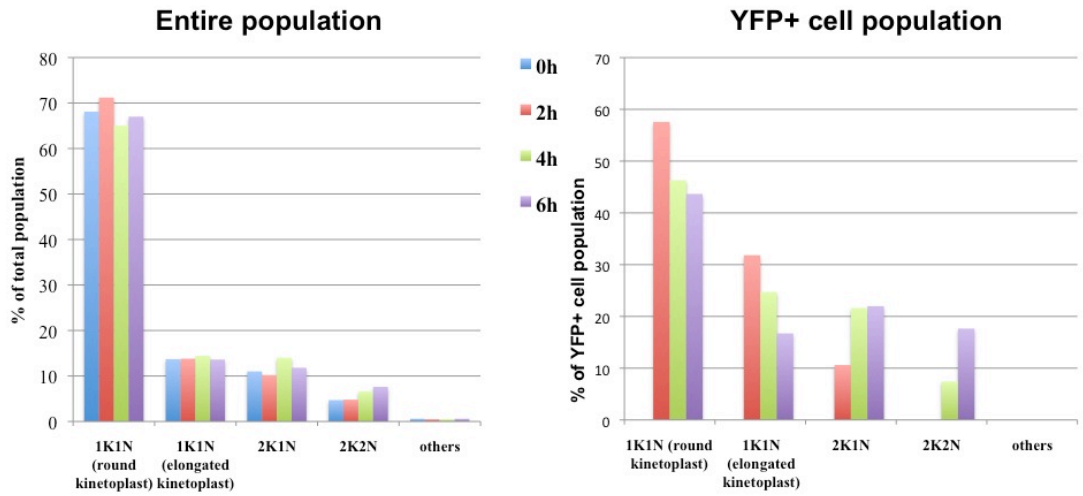


Figure 4.3. Temporal localization of CAP5.5::YFP::Ty following induction. A. Phase contrast and fluorescence micrographs of procyclic cells over-expressing CAP5.5::YFP::Ty after 0, 2, 4 and 6 hours of tetracycline-based induction. The red arrow highlights a 1K1N (elongated K) cell. **B**. Cytoskeletons (detergent-extraction) of cells at the same time points of induction as A were indirectly labelled using anti-GFP antibodies. **C.** Phase contrast and fluorescence micrographs of detergent extracted cells induced for CAP5.5::YFP::Ty (**green**) over-expression for two hours. The monoclonal YL1/2 antibody (**red**) is specific for tyrosinated α-tubulin, a marker for dynamic microtubules. **D**. Graphs show the kinetoplast/nucleus distribution, as assessed by DAPI staining, of cells at 0, 2, 4, 6 hours post-induction. Entire population here denotes all the 500 cells that were counted at each time point, regardless of whether these cells were actually expressing YFP. YFP+ cells represent the proportion of the entire population at these time points that were fluorescent. Others refer to kinetoplast/nuclear configurations other than 1K1N, 2K1N and 2K2N.

and nuclear DNA (Woodward & Gull, 1990). In cells induced for CAP5.5 over-expression, irrespective of whether these cells had YFP signal or not, four types of cells corresponding to 1K1N, 1K1N (elongated K), 2K1N and 2K2N configurations were evident in relative proportions at the various time points (Fig. 4.3D; n=500 at each time point). The relative cell cycle positions of the induced population at these time points did not change during the induction, suggesting minimal perturbations of the cell cycle due to CAP5.5 over-expression.

An elongated or bilobed kinetoplast (K) refers to cells in the kinetoplast division period and in the nuclear G2 phase (Woodward & Gull, 1990). Such cells were interpreted as poised to enter the later periods of the cell cycle as it has been postulated that entry to cytokinesis may be more dependent on kinetoplast division and segregation than on mitosis (Ploubidou *et al.,* 1999). "Others" refer to aberrant cells with kinetoplast/nuclear DNA configurations other than those outlined above. Out of such cells (n=500 at every time point), the DAPI distribution of only those YFP-positive cells following induction was analyzed in Fig 4.5D. This was done to understand whether cells needed to be at a certain stage in the cell division cycle to incorporate CAP5.5 into the cytoskeleton as determined by native fluorescence microscopy. At 2 hours after induction, there was an absence of 2K2N cells (late anaphase) that were YFP-positive, a proxy for CAP5.5::YFP::Ty incorporation into the cytoskeleton. The data suggests that 2K2N cells do not build CAP5.5 into their cytoskeletons, as most of the CAP5.5 incorporation has already happened in the earlier stages of the cell cycle. Instead, cells with posterior-end restricted CAP5.5 expression were mostly at an earlier stage [1K1N; 1K1N (elongated K)] of the cell cycle at 2 hours post-induction. At later time points,

YFP+ cells at later stages of the cycle (2K1N, 2K2N) increased with a corresponding decrease in cells with 1K1N and 1K1N (elongated K) configurations. This is attributable to the progression of the cells through the cell division cycle, where cells that had built CAP5.5 into their cytoskeletons at earlier stages of the cell cycle have transitioned into the later phases of the cell cycle.

Analysis of the current data suggests a possible cell cycle regulation of CAP5.5 incorporation into the cytoskeleton, with CAP5.5 being added to the sub-pellicular cytoskeleton from the posterior end of the cell at earlier stages of the cell cycle. This suggests an elegant mechanism for cell cycle regulated incorporation of MAPs like CAP5.5 into the cytoskeleton. By inducibly expressing CAP5.5 using an ectopic 3' UTR, one is able to dissociate the expression of CAP5.5 protein from its incorporation into the cytoskeleton. Therefore, although CAP5.5::YFP::Ty expression is expected to be independent of the cell cycle in our system, its incorporation into the sub-pellicular cytoskeleton is not.

4.5 Characterization of CAP5.5 N6::YFP::Ty

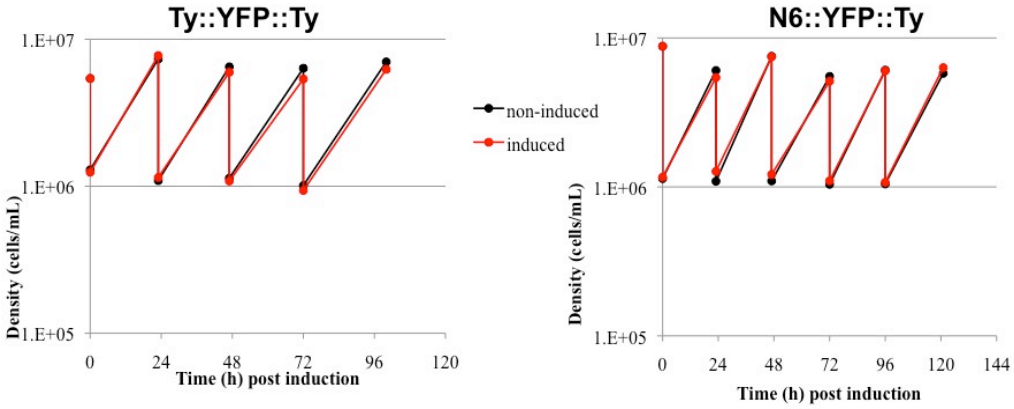
I was also interested in investigating the molecular determinants in CAP5.5's sequence that are involved in targeting it to the cell body cytoskeleton, when the default pathway for dually acylated proteins, such as CAP5.5, is to the flagellum (Emmer *et al.,* 2009; Liu *et al.,* 2010). *In vivo* labelling of CAP5.5 with radioactive myristic acid and palmitic acid has confirmed the presence of these fatty acid modifications (Hertz-Fowler *et al.,* 2001). CAP5.5 and CAP5.5V contain a myristolyation motif (MGXXXS/T),

corresponding to its first 6 amino acids MGCGGS (Towler *et al.,* 1987). N-myristoylation is a co-translational process in which protein sequences with such a motif have myristic acid added to the glycine by an amide bond, following the removal of the N-terminal methionine (Resh, 2006). Although there is no consensus sequence for palmitoylation, for dually acylated proteins containing the sequence Met-Gly-Cys at their N-terminus, Gly2 myristolyation facilitates the thioester bond linkage of palmitic acid to the downstream cysteine (Resh, 1999; Godsel & Engman, 1999; Mumby, 1994). Protein lipidation can regulate membrane targeting and trafficking (Resh, 2006).

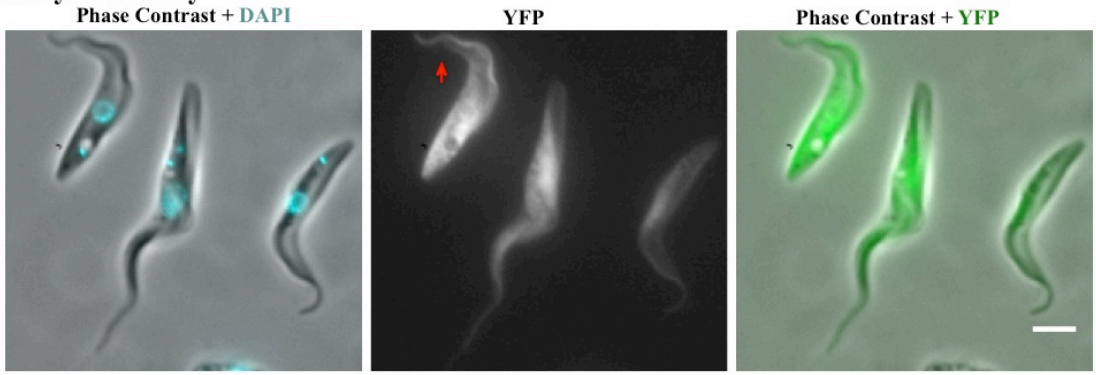
Dually acylated proteins such as phosphoinositide-specific phospholipase C (Pl-PLC) and flagellar calcium binding protein (FCaBP) in *T. cruzi,* for example, contain a short stretch of amino acids in their N-terminus (< 25aa) that contain the necessary signals to target the two proteins to the plasma and flagellar membranes respectively (Martins *et al.,* 2010; Maric *et al.,* 2011). For a protein like CAP5.5 that is strongly associated with the sub-pellicular cytoskeleton as well as the cell body plasma membrane, several YFP-containing gene fusions were generated for over-expression in procyclic *T. brucei* ranging from the first 6 amino acids to domains I-III (containing domains I, II, and III) with the aim of studying motifs or domains important in CAP5.5's localization.

Over-expression of MGCGGS (N6) fused to a C-terminal YFP::Ty tag using the pDex577 vector in procyclic parasites did not exhibit a noticeable growth defect over a period of five days when compared to non-induced populations (Fig. 4.4A). Ty::YFP::Ty alone was expressed as a control and exhibited no growth defect upon induction either (Fig. 4.4A). After 24 hours of induction, both the Ty::YFP::Ty and N6::YFP::Ty fusion

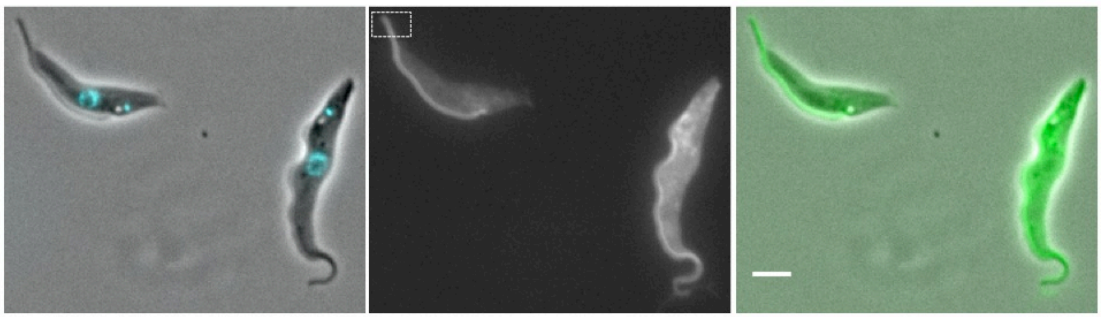
A



B. Ty::YFP::Ty



N6::YFP::Ty



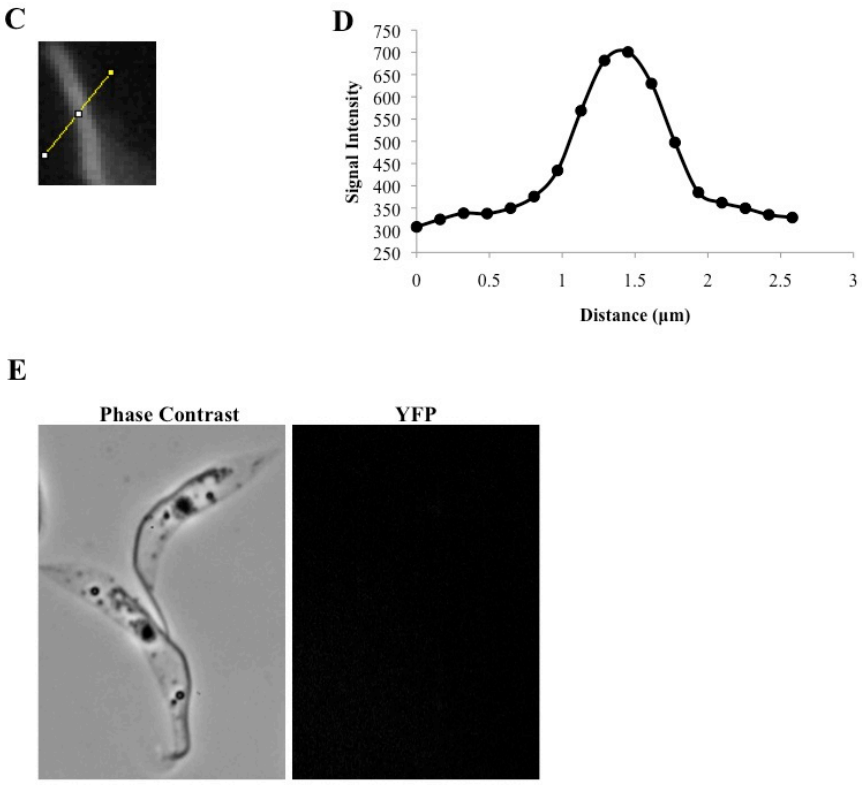


Figure 4.4. N6::YFP::Ty localizes to the cell body and flagellum. A. Growth curve of induced and non-induced cells over-expressing either Ty::YFP::Ty or CAP5.5's N6::YFP::Ty in procyclic cells using the pDex577 vector. **B**. Phase contrast and fluorescence micrographs of 2% PFA-fixed of cells over-expressing YFP::Ty or N6::YFP::Ty after 24 hours of induction. Red arrows refer to the cell body and flagellum junction to highlight that the labelling continues beyond the cell body and into the flagellum. DNA was labelled with DAPI (cyan). **D.** The distribution of YFP signal intensity along the yellow line (**C**) drawn across the flagellum of the induced N6::YFP::Ty cell in B. **E**. Cells from the same induction as B were treated with detergent (cytoskeleton), fixed in methanol and their localizations assessed by native fluorescence microscopy. Scale bar, 3 µm.

Proteins localized both to the cell body and the flagellum, and did not associate strongly to the cytoskeleton upon detergent extraction (Fig. 4.4B; 4.4E). The addition of only 6 amino acids to YFP::Ty as in the case of N6::YFP::Ty showed stronger flagellar labelling relative to the cell body, when compared to the Ty::YFP::Ty control.

The YFP fluorescence signal intensity distribution along the flagellum had a normal distribution (Fig. 4.4C, D), which suggested localization in the flagellar matrix. Had the localization been in the flagellar membrane, one would have expected a biomodal distribution profile. However, another possibility might be YFP labelling existed both in the flagellar matrix and the flagellum membrane, hence not giving a detectable bimodal distribution.

The localization observed is consistent with the observation of a cytosolic localization when N6::GFP was over-expressed in *T. cruzi* epimastigotes (Martins *et al.,* 2010). This suggests that while lipidation is known to be essential for membrane association, dual acylation alone is not sufficient and requires additional trafficking elements. No metabolic labelling of N6::YFP::Ty was done to confirm dual acylation *in vivo*, however, others have shown that the 6 amino acid-long acylation motif can be dually acylated when fused to GFP and expressed in *T. cruzi* (Martins *et al.,* 2010).

Therefore, it is possible that the predominant flagellar labelling seen in N6::YFP::Ty relative to Ty::YFP::Ty may be due to the affinity of dually acylated YFP to the flagellum membrane where there is an enrichment of sterols, glycolipids, sphingolipids, all of which are components of canonical lipid rafts in contrast to the microenvironment of the cell body (Chailley, 1985; Tetley, 1986; Kaneshiro *et al.,* 1984).

Association of dually acylated proteins with such lipid rafts has been suggested as a possible mechanism for flagellar membrane trafficking in kinetoplastids (Fridberg *et al.,* 2006; Tyler *et al.,* 2009). However, dual acylation alone, i.e. the six amino acid motif, is not sufficient for strong flagellum membrane attachment and requires additional signalling elements to stabilize such interactions.

4.6 Characterization of CAP5.5 N37::YFP::Ty

The N-terminus of CAP5.5 contains a proline-rich region immediately following the myristoylation-palmitoylation signal. The sequence PxxP conforms to the binding sequence motif found in many ligands of SH3- type domains (Ren *et al.,* 1993; Yu *et al.,* 1994). Proline-rich regions are recognized as mediating protein-protein interactions, particularly at the plasma membrane interface (Buday *et al.,* 1999; Niebuhr *et al.,* 1997). However, this stretch of proline-rich sequence -comprising the first 37 amino acids of CAP5.5 (N37)- remains uncharacterized in *T. brucei.* Additionally, for many dually acylated proteins like Pl-PLC and FCaBP in *T.cruzi,* a short stretch of the N-terminus (<25 aa) is sufficient to direct these proteins to their distinct localizations (for a review see Fridberg *et al.,* 2006). Therefore, it is possible that the pellicular localization signal for dually acylated CAP5.5 is also within a similar stretch of amino acids.

Over-expression of N37::YFP::Ty (pDex577) in procyclic cells was validated by western blotting using BB2, an anti-Ty antibody (Fig. 4.5A; Bastin *et al.,* 1996). Two bands with molecular weights between 30 and 35 kDa were also observed in addition to that expected of the fusion protein (29 kDa; Fig. 4.5A). These two extra bands were observed in two separate trials of induction and may correspond to possible post-

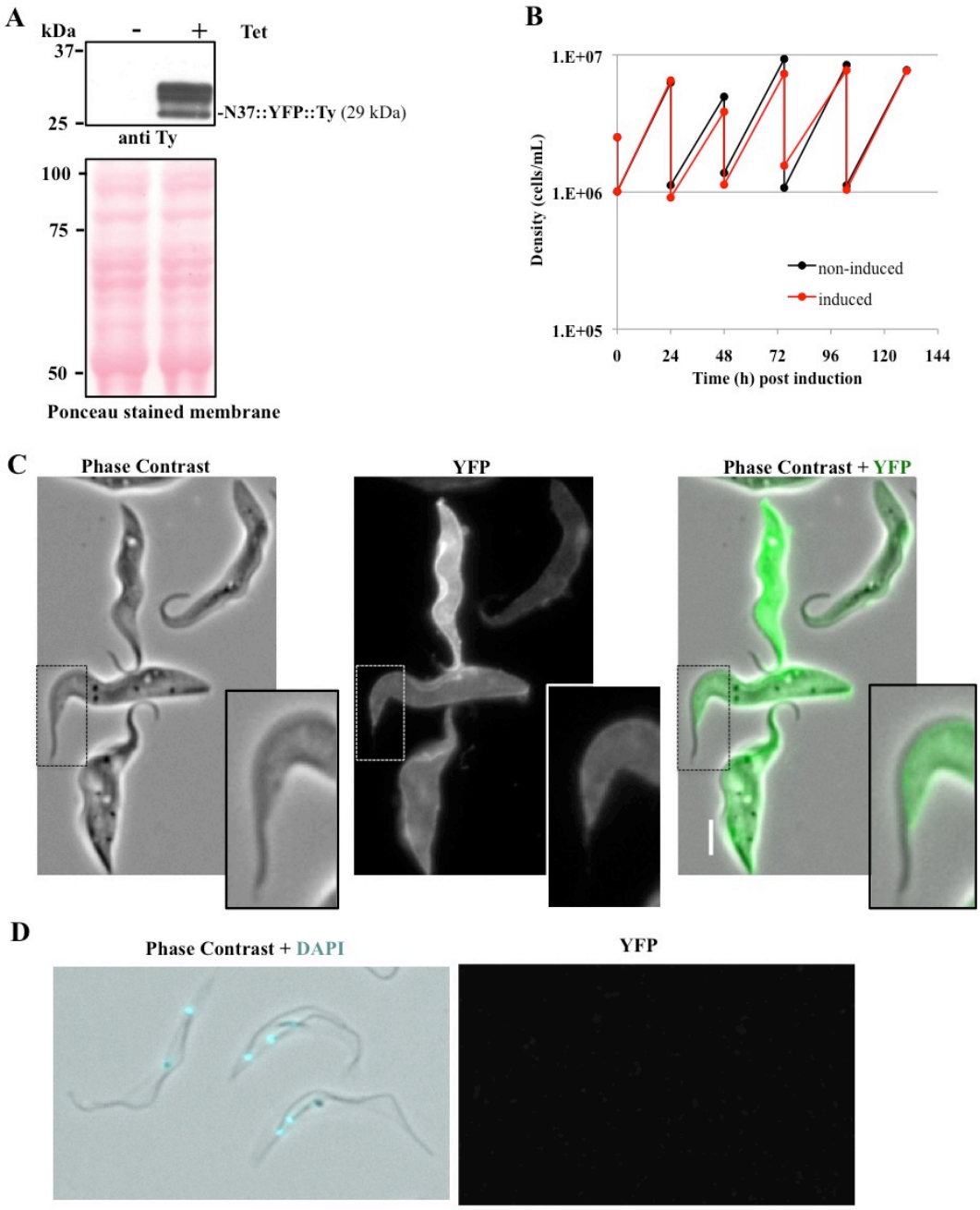


Figure 4.5. N37::YFP::Ty is restricted to the cell body. A. Western blot of 3.5 x10^6 cells/lane of cells over-expressing CAP5.5's N37::YFP::Ty (pDex577 vector) under tetracycline induced (+) or non-induced (-) conditions using BB2, an anti-Ty epitope antibody (Bastin *et al.,* 1996). Ponceau-stained membrane is for loading control. **B**. Growth curve of induced and non-induced cells. **C**. Phase and fluorescence micrographs of 2% PFA-fixed of cells over-expressing N37::YFP::Ty after 24 hours of induction. **D**. Cells from the same induction as C were treated with detergent (cytoskeleton), fixed in methanol and their localizations assessed by native fluorescence microscopy. Scale bar, 3 µm.

translational modifications. Growth curve analyses of non-induced and cells induced for expression of N37::YFP::Ty did not reveal any clear growth defect (Fig. 4.5B). N37::YFP::Ty localized predominantly to the cell body and appeared to be excluded from the flagellum in contrast to N6::YFP::Ty (Fig. 4.7C). Additionally, an accumulation in the cell surface indicative of membrane-association was also observed.

While N37::YFP::Ty gives a distinct pellicular membrane signal, N6::YFP::Ty does not but does have a distinct flagellum signal. Upon non-ionic detergent extraction, cells over-expressing N37::YFP::Ty did not have any YFP staining (Fig.4.7D). Taken together, this suggests that N37::YFP::Ty localizes to the pellicular membrane but is not associated with the cytoskeleton. The difference in localization between N37::YFP::Ty and N6::YFP::Ty highly suggests that the addition of the extra 29 amino acids changes the interaction of the fusion protein with the lipid environment of the flagellum, such that it is no longer retained in the flagellum, explaining the primary localization of N37::YFP::Ty to the pellicular membrane. In *T. cruzi*, conserved lysines in the first 13-24 amino acids of FCaBP were shown to confer localization specificity and lipid raft affinity (Maric *et al.,* 2011). The possibility of additional trafficking elements, such as one based on charge distribution like FCaBP, in the N-terminus of CAP5.5 that promotes its retention in the pellicular membrane and the temporal order of such an event needs to be seriously considered in the future.

4.7 Characterization of CAP5.5 domain I::YFP::Ty

Expression of CAP5.5 domain I::YFP::Ty was confirmed by western blotting using an anti-TY monoclonal antibody (Fig. 4.6A; Bastin *et al.,* 1996). There was no

measurable growth defect in clones over-expressing CAP5.5 domain I::YFP::Ty, irrespective of tetracycline induction (Fig. 4.6B). To determine the cellular localization of domain I::YFP::Ty, transfectants expressing this fusion protein were analyzed by native fluorescence microscopy. This showed that YFP-tagged domain I localized to the cell body membrane and not the flagellum, and such a labelling did not survive detergent extraction, showing that it is not strongly associated with the cytoskeleton (Fig. 4.6C-D). This suggests that domain I is not be the domain mediating association to the sub-pellicular cytoskeleton. Although the function of CAP5.5's domain I is unknown, it contains the dual acylation motif and a proline-rich motif which are thought to mediate CAP5.5's interactions with the cytosolic face of the plasma membrane.

4.8 Localization of domain I-II::YFP::Ty

Domain II has strong homology to the protease core domain of typical calpains but has two distinct indels (see chapter 3). Recombinant CAP5.5V domain II had no detectable protease activity *in vitro* and the function of this conserved domain in kinetoplastid calpains remains hitherto uncharacterized. We wanted to determine if this domain is involved in CAP5.5's trafficking by over-expressing a fusion construct encoding domains I and II (DI-DII) with a C-terminal YFP::Ty tag. Ectopic expression of DI-II::YFP::Ty did not have a noticeable growth phenotype (Fig. 4.7A). Localization analyses of such cell lines by native fluorescence microscopy after 24 hours of induction revealed approximately 3 out of every 500 cells that were YFP positive (Fig. 4.7B). The labelling appears largely in the cell body although signal is visible in the flagellum. Such labelling looks quite similar to that of cells expressing Ty::YFP::Ty alone.

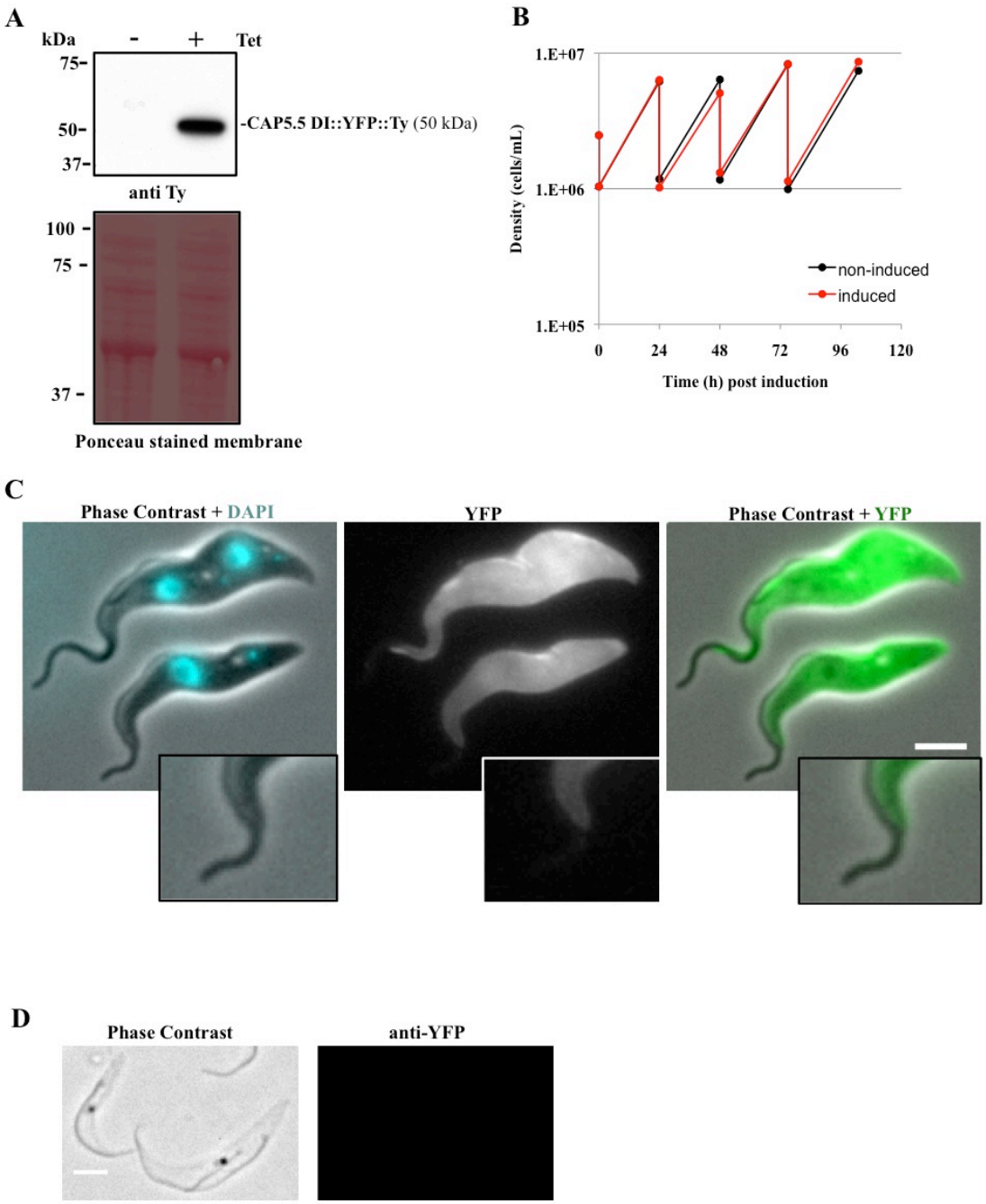


Figure 4.6. Ectopic expression of DI::YFP::Ty in procyclic cells. A. Western blot of cells over-expressing DI::YFP::Ty using anti-Ty antibody. Ponceau-stained membrane is for loading control. **B**. Growth curve of induced and non-induced cells. **C**. Phase contrast and fluorescence micrographs of 2% PFA-fixed of cells over-expressing DI::YFP::Ty after 24 hours of induction. **D**. Cells from the same induction as C were treated with detergent (cytoskeleton), fixed in methanol and their localizations assessed by immunofluorescence microscopy with an anti-GFP antibody. Scale bar, 3 μm.

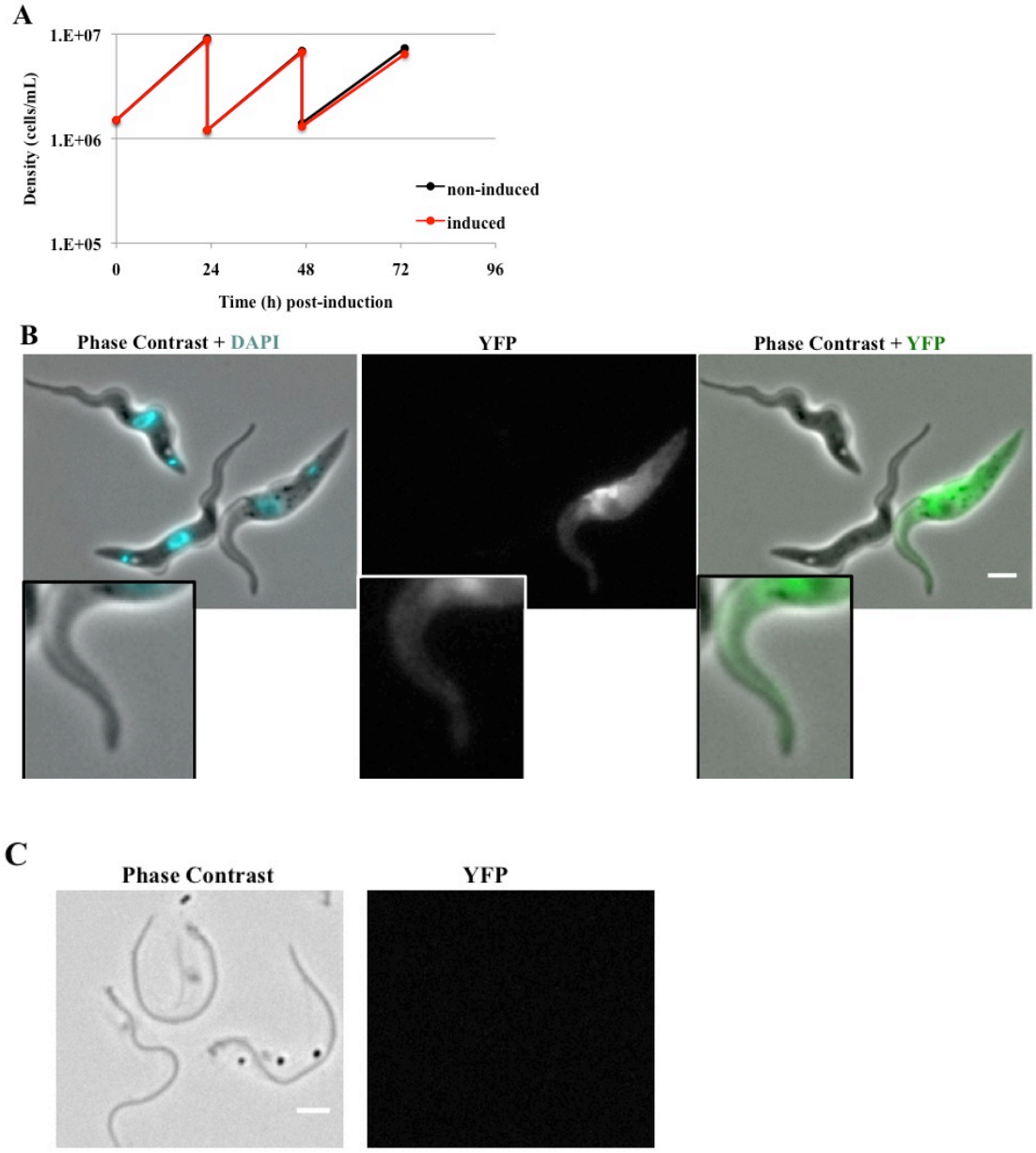
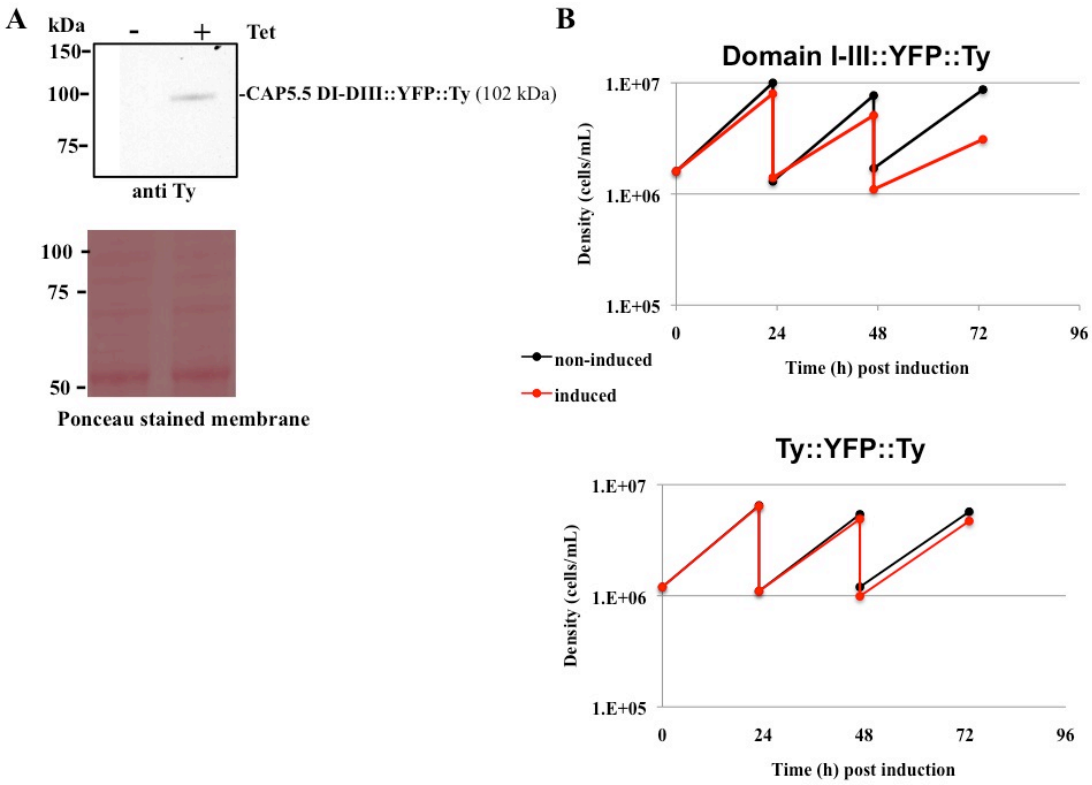


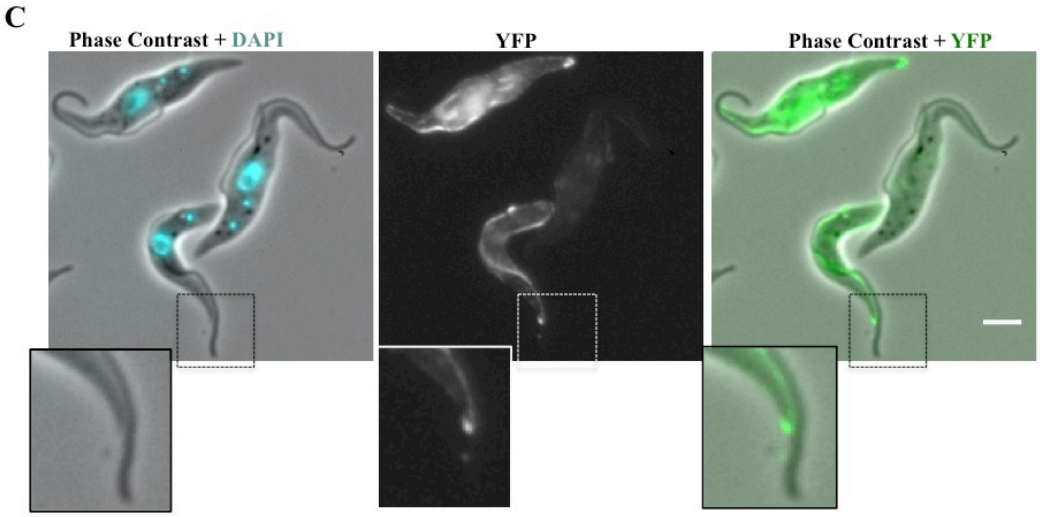
Figure 4.7. Ectopic expression of CAP5.5 domains I-II. A. Growth curve of cells either induced or not for ectopic expression of DI-DII::YFP::Ty. **B**. Phase contrast and fluorescence micrographs of cells induced for 24 hours. **C**. Micrographs of cells at the same stage of induction as B after detergent extraction (cytoskeleton). Scale bar, 3 µm. Work encompassed in this figure was contributed by Samuel Dean apart from the cloning of the construct.

There is no edge effect suggesting the fusion protein is not associated with the pellicular membrane, however, there is a concentrated signal at the mid-ventral position (Fig. 4.7B). Such a signal may possibly be in the endoplasmic reticulum and therefore, an indication of a badly processed protein, which would explain the small number of YFP positive in the induced population. Such labelling did not survive detergent extraction, arguing that the fusion protein is not cytoskeleton-associated (Fig. 4.7D). It is possible that DI-II::YFP::Ty, when expressed alone, may be misfolded and not processed properly, thereby targeted for degradation by the cellular machinery.

4.9 Expression and localization of domain I-III::YFP::Ty

The last in the series of expression constructs designed to study the molecular determinants of CAP5.5's trafficking to the sub-pellicular cytoskeleton was CAP5.5's domain I, II and III (Δ domain IV). Its over-expression (pDex577) in procyclic cells was validated by western blotting using anti-Ty antibodies, and a band corresponding to the expected size of the fusion protein (102 kDa) was observed (Fig. 4.8A). Over-expression of CAP5.5 domain I-III::YFP::Ty, but not Ty::YFP::Ty, yielded a growth defect phenotype by 48 hours post tetracycline induction (Fig. 4.8B). Calculation of the doubling times revealed that in the first 23 hours of induction, the induced population had a doubling time of 9.205 h compared to 8.699 h in uninduced populations. By 47 hours, the induced population's doubling time was 12.868 h compared to 9.350 h in uninduced; and the doubling time was 17.394 h and 11.039 h in induced and uninduced populations at 73h respectively. This is in contrast to the absence of growth phenotypes when over-expressing the full-length protein in the same vector (Fig. 4.2). While it is likely that the


A
kDa
-
+
Tet
150
100
75
-CAP5.5 DI-DIII::YFP::Ty (102 kDa)
anti Ty
100
75
50
Ponceau stained membrane
B
Domain I-III::YFP::Ty
1.E+07
1.E+06
1.E+05
Density (cells/mL)
0
24
48
72
96
Time (h) post induction
non-induced
induced
Ty::YFP::Ty
1.E+07
1.E+06
1.E+05
Density (cells/mL)
0
24
48
72
96
Time (h) post induction


C
Phase Contrast + DAPI
YFP
Phase Contrast + YFP

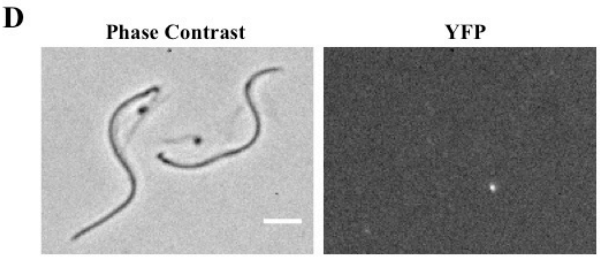

D
Phase Contrast
YFP

Figure 4.8. Ectopic expression of domains I-III results in a growth defect phenotype. **A.** Western blot of 3.5E6 cells/lane of cells over-expressing CAP5.5's DI-DIII::YFP::Ty (pDex577 vector) under tetracycline induced (+) or non-induced (-) conditions using anti-TY epitope antibody. Ponceau-stained membrane is for loading control. **B**. Growth curves of non-induced cells and those induced for ectopic expression of Ty::YFP::Ty and DI-DIII::YFP::Ty using the same vector. The doubling times corresponding to this graph are included in the text. **C**. Phase contrast and fluorescence micrographs of 2% PFA-fixed of cells over-expressing DI-DIII::YFP::Ty after 24 hours of induction. DNA was labelled with DAPI (cyan). **D.** Cells from the same induction as **C** were treated with detergent (cytoskeleton), fixed in methanol and their localizations assessed by native fluorescence microscopy. Scale bar, 3 μm. Figures 4.8B-4.8D were contributed by Samuel Dean.

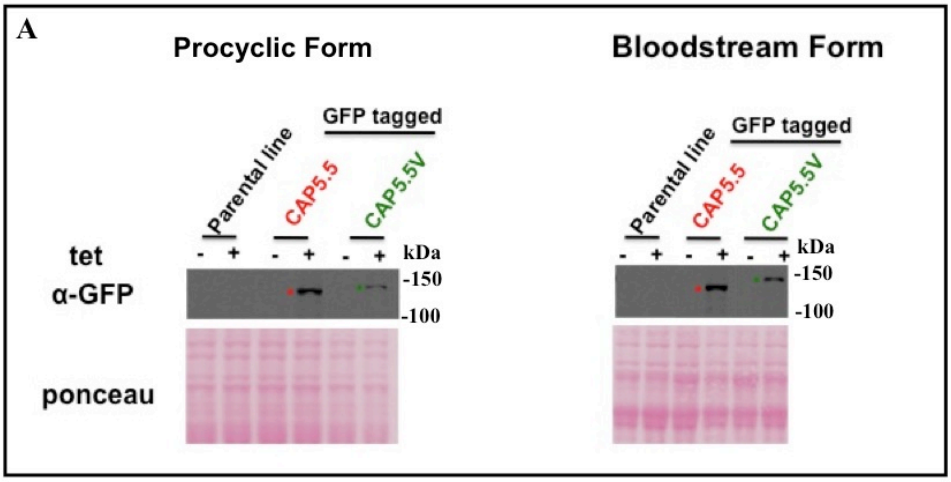
ectopic full-length CAP5.5 at the present levels simply performs the function of native CAP5.5, with DI-DIII, the truncation of domain IV may exert a dominant biological effect. Such an effect may be due to the relief of regulation by domain IV on CAP5.5's activity or the cellular response to an improperly folded protein resulting in toxicity. Localization analyses by native fluorescence microscopy of induced cells showed a clear labelling around the surface of the cell, suggesting membrane association (Fig. 4.8C). However, the pellicular signal was not homogenous or uniform but in all YFP-positive cells, there was no clear labelling in the flagellum. Additionally, some cells were labelled with a spot at the anterior tips of their cell bodies (~19%). All labelling were susceptible to detergent extraction indicating that the fusion protein is not strongly associated to the cytoskeleton (Fig. 4.8D). Taken together, this data suggests that since there was no labelling in the cytoskeleton preparations of cells over-expressing CAP5.5 domains I-III albeit with a C-terminal YFP::Ty, that the excluded domain IV may be the domain mediating cytoskeletal interaction as CAP5.5 has been shown to bind to the cytoskeleton by both biochemical and structural analyses (Hertz-Fowler *et al.,* 2001).

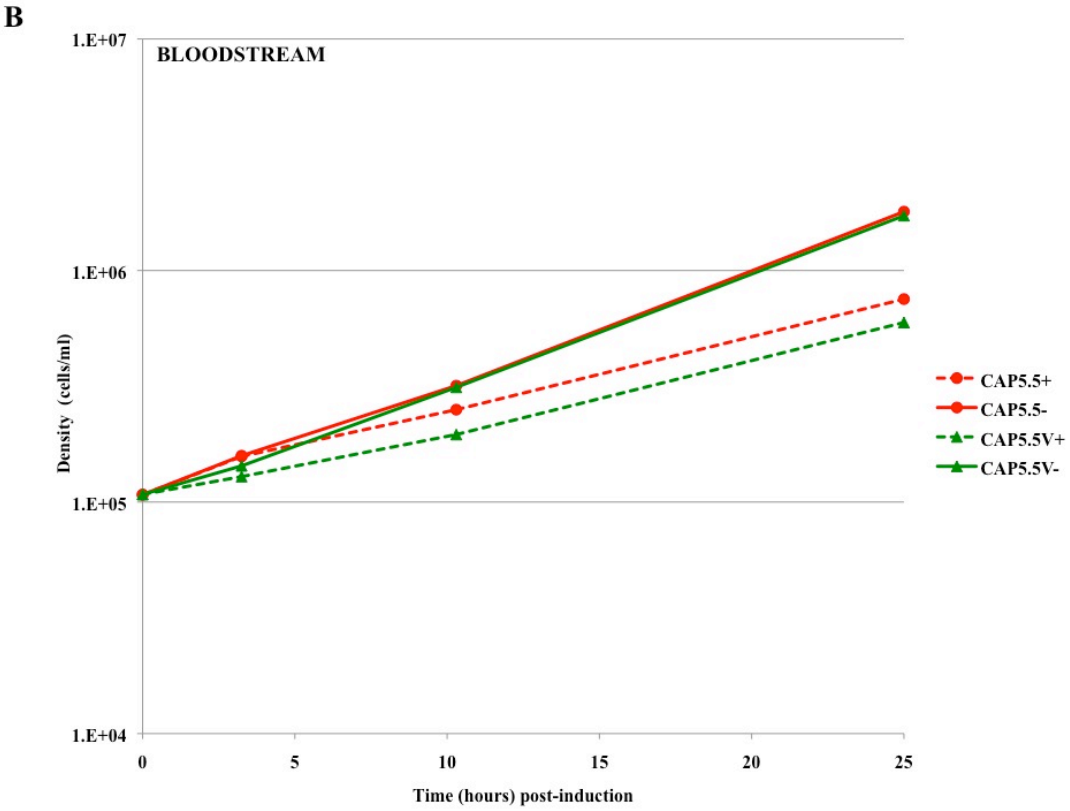
4.10 Characterization of GFP::CAP5.5 and GFP::CAP5.5V

Two constructs that over-express an N-terminally GFP tagged CAP5.5 and CAP5.5V were also engineered into pDex777 vectors, which utilize the PFR-2 3'UTR and therefore, has a higher expression level than when pDex577 vector is used (Poon *et al.,* 2012). These constructs were transfected into both procyclic and bloodstream parasites. The positioning of around 238 amino acids (~27 kDa) N-terminal to the starting methionine of both CAP5.5 and CAP5.5V is hypothesized to sterically (and otherwise)

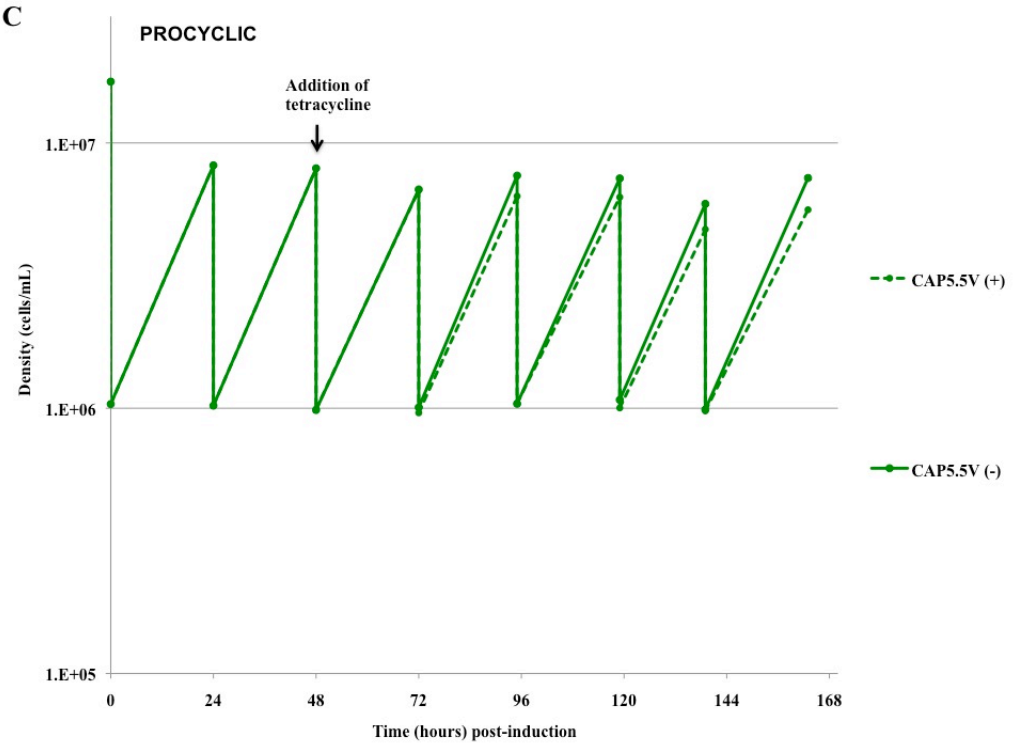
hinder any N-myristoyltransferases that add a myristic acid to the glycine by an amide bond, following the removal of the N-terminal methionine (Resh, 2006). GFP does not contain the myristoylation/palmitoylation motif. Indeed, when an N-terminal GFP or HA tag was fused to small myristoylated proteins (SMP) in *Leishmania* spp., N-myristoylation (and additional acylation modifications) was disrupted along with the normal cellular trafficking routes (Malcolm McConville, personal communication). SMPs are either monoacylated (myristic acid) or diacylated (myristic and palmitic acids), which have strong sequence and structural homology to the domain I of CAP5.5/CAP5.5V (Tull *et al.,* 2004; see Section 3.2). Therefore, over-expressing CAP5.5/CAP5.5V with an N-terminal tag is likely to not be acylated, although biochemical labelling *in vivo* was not performed.

Ectopic expression of GFP tagged CAP5.5 and CAP5.5V in both procyclic and bloodstream forms of *T. brucei* was validated by Western blotting using anti-GFP antibodies (Fig 4.9A). Cell densities of parasites ectopically expressing CAP5.5 or CAP5.5V were monitored in order to determine the effect of such expression on growth rates. In bloodstream forms, pronounced growth defects were observed within the first 10 hours of induction of cells ectopically over-expressing either GFP::CAP5.5 or GFP::CAP5.5V (Fig. 4.9B). In procyclic cells, a mild growth defect was observed when over-expressing either fusion protein, albeit after 24 hours of induction for the procyclic-specific CAP5.5 and 48 hours for the bloodstream specific CAP5.5V (Fig. 4.9C). Procyclic cells expressing GFP:: CAP5.5 had a difference in doubling time of ~1.01 h (9.233 h- 8.216 h) upon induction within 24 hours of induction; a difference of 0.932 h (8.918 h-7.987 h) and 0.652 h (8.818 h- 8.165 h) in 48 hours and 72 hours of induction.


A
Procyclic Form
Bloodstream Form
GFP tagged
Parental line
CAP5.5
CAP5.5V
tet
α-GFP
ponceau
kDa
-150
-100
- + - + - +


B
BLOODSTREAM
1.E+07
1.E+06
1.E+05
1.E+04
Density (cells/ml)
0
5
10
15
20
25
Time (hours) post-induction
CAP5.5+
CAP5.5-
CAP5.5V+
CAP5.5V-

C


PROCYCLIC
Addition of tetracycline
1.E+07
1.E+06
1.E+05
Density (cells/mL)
0
24
48
72
96
120
144
168
Time (hours) post-induction
CAP5.5V (+)
CAP5.5V (-)

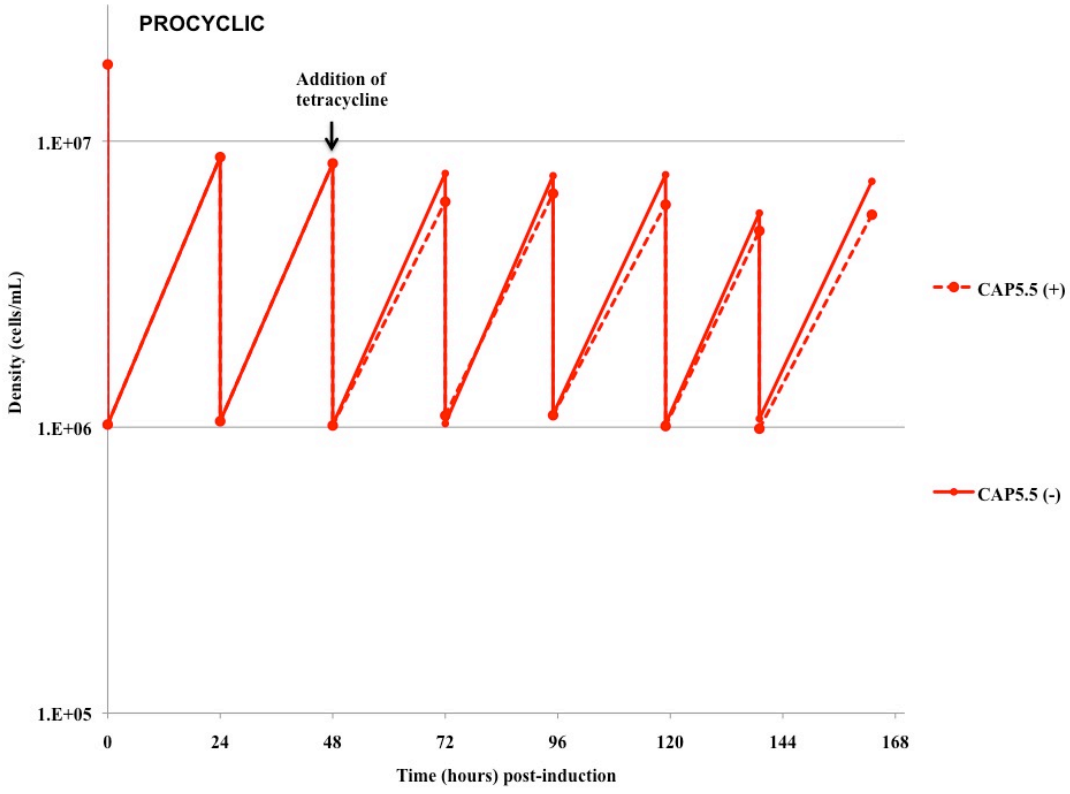

PROCYCLIC
Addition of tetracycline
1.E+07
1.E+06
1.E+05
Density (cells/mL)
0
24
48
72
96
120
144
168
Time (hours) post-induction
CAP5.5 (+)
CAP5.5 (-)

Figure 4.9. Ectopic expression of GFP::CAP5.5 and GFP::CAP5.5V lead to growth defects. A. Western blot analysis of *T. brucei* procyclic forms and bloodstream forms transfected with pDex777 constructs to over-express N-terminally GFP tagged CAP5.5 and CAP5.5V. Blots containing whole-cell protein lysates induced in the presence (+) or absence (-) of tetracycline for 24 hours were probed with antibodies against GFP. **B**. Cell densities of bloodstream parasites over-expressing CAP5.5/CAP5.5V in a 24 hour time period of induction. **C.** Growth curves of procyclic parasites ectopically expressing GFP:: CAP5.5/CAP5.5V under either induced (+) or non-induced (-) conditions. The arrow marks the point when tetracycline was first added to the cultures. Doubling times corresponding to this graph are presented within the text. Bloodstream and procyclic cells over-expressing GFP (pDex777 vector) did not have any growth defects (data not shown).

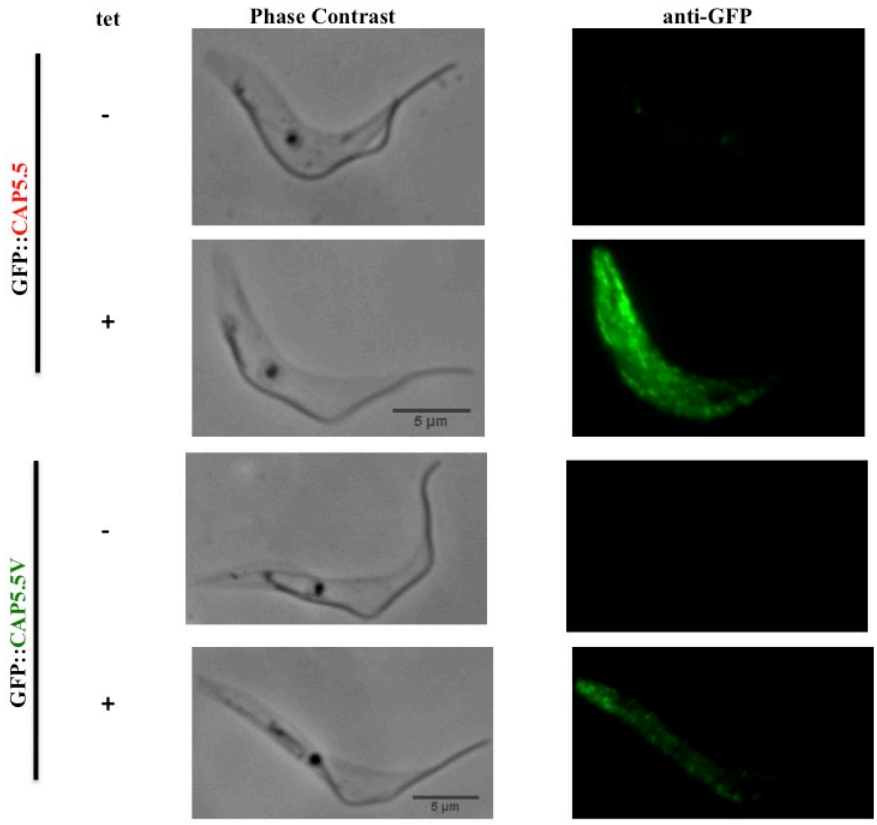
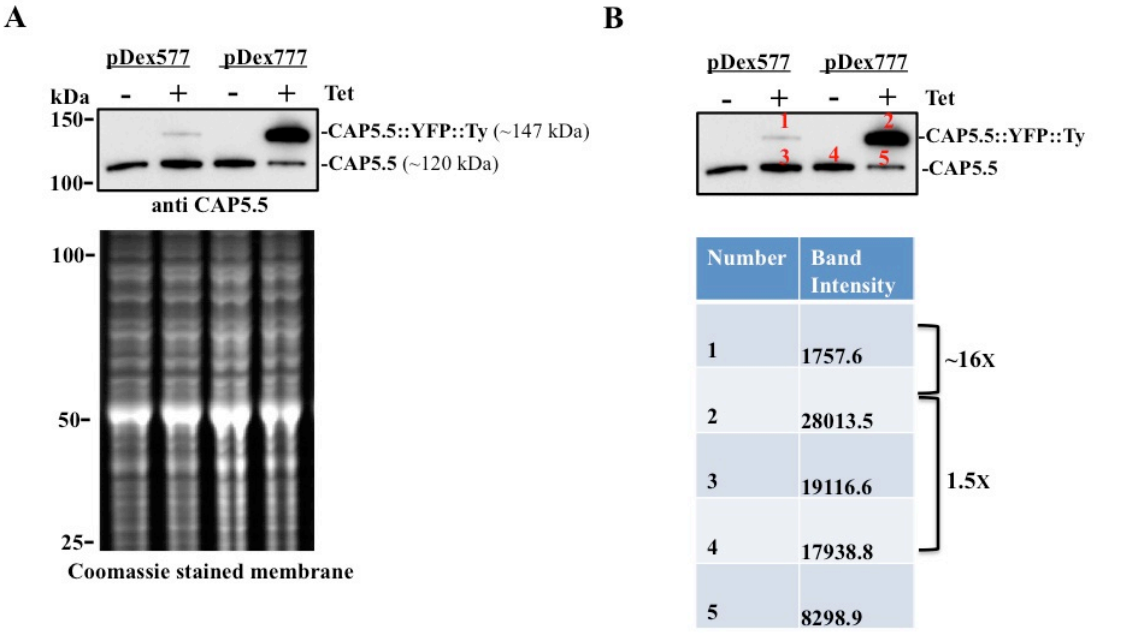
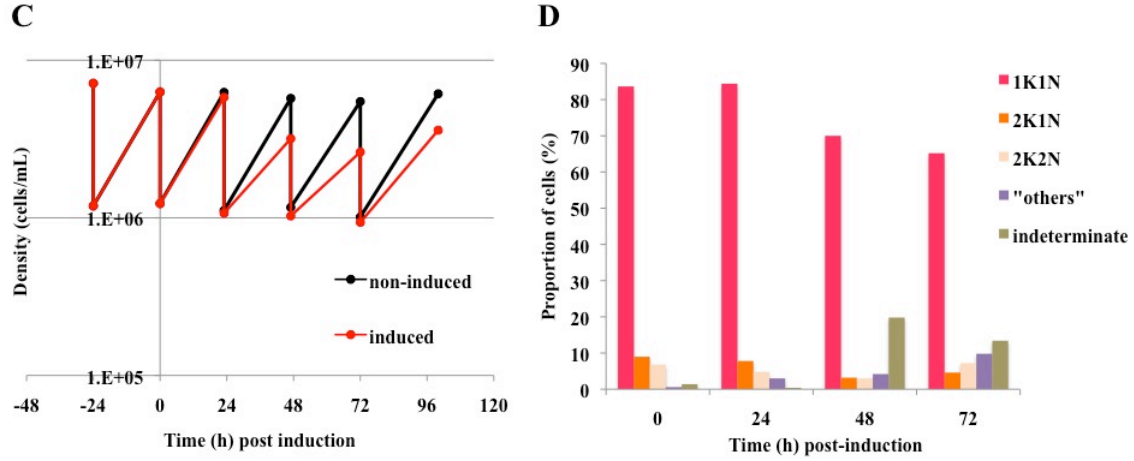


Figure 4.10. GFP::CAP5.5V localizes to the cytoskeleton of procyclic cells. Phase contrast and indirect immunofluorescence micrographs (using anti-GFP antibodies) of detergent extracted and methanol fixed cells that had been induced to ectopically express GFP::CAP5.5/CAP5.5V over a period of 24 hours of induction (+, induced; -, non-induced). Scale bar, 5 µm.

Procyclic cells over-expressing GFP::CAP5.5V had differences in generation time of 0.056 h (8.746-8.690), 0.552 h (8.477-7.924) and 0.794 h (8.945-8.151) in 24, 48 and 72 hours post induction. Ectopic expression of GFP in either lifecycle stages had no effect on their respective growth (data not shown). Attachment of lipophilic moieties to proteins facilitates protein interactions, membrane binding and targeting, and intracellular signalling (Resh, 1996). Besides the established roles of the cytoskeleton in determining cell shape and motility, it can also be the platform for intracellular signal transduction by linking cytoskeletal proteins to the membrane and membrane receptor proteins (Schoenwaelder & Burridge, 1999; Barth *et al.,* 1997). It is therefore possible that when CAP5.5/CAP5.5V are not acylated, the normal functions that the acylation confers on these proteins are also disrupted suggesting a possible function for CAP5.5/CAP5.5V in protein-protein interactions or intracellular signalling.

Procyclic cells ectopically expressing GFP::CAP5.5 and GFP::CAP5.5V were detergent extracted (cytoskeletons) and their cellular localizations determined by indirect immunofluorescence microscopy using anti-GFP antibodies (Fig. 4.10). One notices that the procyclic-specific GFP::CAP5.5 localized to the procyclic cytoskeleton but so does the bloodstream-specific GFP::CAP5.5V (Fig 4.10). The difference in the global labelling of the two fusion proteins in procyclic cells may be attributable to the difference in expression levels as indicated by western blot. There are no major ultrastructural differences between the sub-pellicular cytoskeletons of bloodstream and procyclic forms but CAP5.5 and CAP5.5V are a pair of differentially regulated paralogs that perform analogous roles. Here we show that while they may have life-cycle specific expression


A
pDex577
pDex777
kDa
-
+
-
+
Tet
150-
100-
-CAP5.5::YFP::Ty (~147 kDa)
-CAP5.5 (~120 kDa)
anti CAP5.5
100-
50-
25-
Coomassie stained membrane
B
pDex577
pDex777
-
+
-
+
Tet
1
2
3
4
5
-CAP5.5::YFP::Ty
-CAP5.5
Number
Band Intensity
1
1757.6
2
28013.5
3
19116.6
4
17938.8
5
8298.9
~16X
1.5X


C
1.E+07
1.E+06
1.E+05
Density (cells/mL)
non-induced
induced
-48
-24
0
24
48
72
96
120
Time (h) post induction
D
Proportion of cells (%)
90
80
70
60
50
40
30
20
10
0
1K1N
2K1N
2K2N
"others"
indeterminate
0
24
48
72
Time (h) post-induction

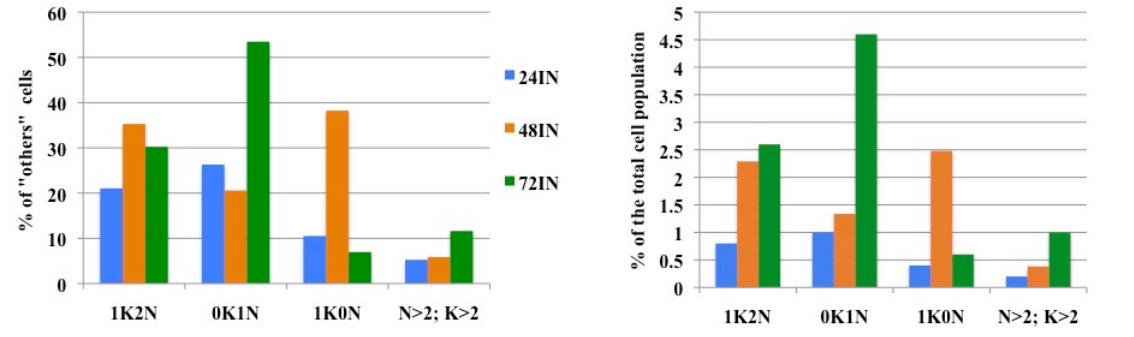
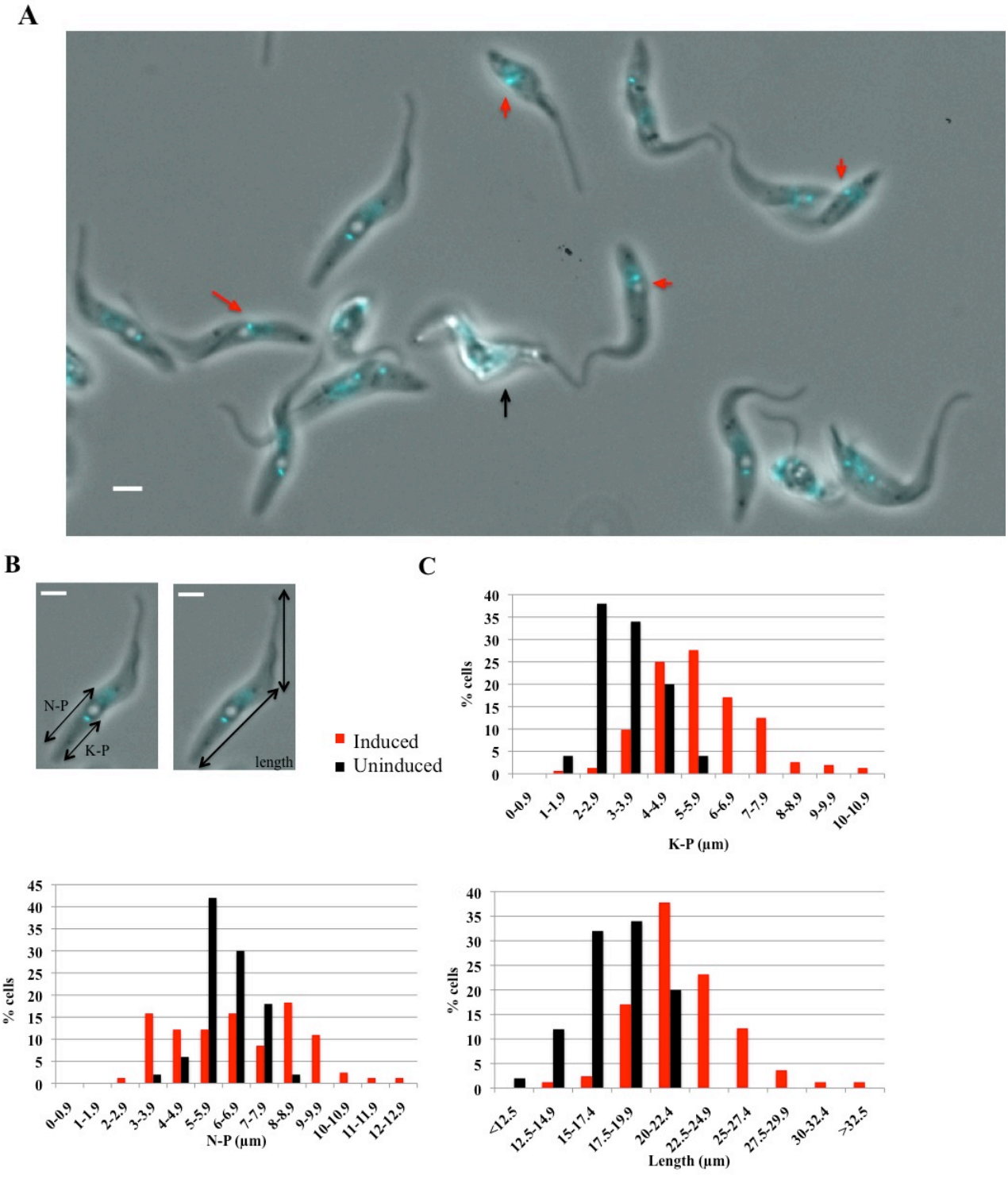

E
% of "others" cells
60
50
40
30
20
10
0
24IN
48IN
72IN
1K2N
0K1N
1K0N
N>2; K>2
% of the total cell population
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
1K2N
0K1N
1K0N
N>2; K>2

Figure 4.11. Over-expression of CAP5.5 in procyclic parasites leads to abnormal cell types. A. Western blot analysis of *T. brucei* procyclic forms transfected with pDex577 (see Chapter 4.2) and pDex777 constructs to over-express C-terminally YFP tagged CAP5.5. Blots containing whole-cell protein lysates induced in the presence (+) or absence (-) of tetracycline for 24 hours were probed with monoclonal antibodies against CAP5.5. A coomassie-stained protein gel is shown as a loading control. **B**. The western blot bands were quantified individually by multiplying the average intensity of all pixels in equally sized regions and the area of such region. The average intensity of the bands were adjusted by subtracting the intensity of the background. All analyses were done using ImageJ software. **C**. Growth curves of induced and non-induced cells. **D**. DAPI analyses of induced cells at 0, 24, 48 and 72 hours of induction (n=500 at each time point). **E**. The left graph of the "others", essentially cells that were not 1K1N, 2K1N or 2K2N, shows the proportions of each category of "others" cells in the entire "others" population at 24, 48 and 72 hours post-induction. The right graph shows 1K2N, 0K1N, 1K0N and multinucleate and/or multi-kinetoplast cells proportionate to the entire cell population at three time points following induction. IN, induced; indeterminate, cells with abnormal morphology that hazes the interpretation of the DAPI labeling.

and roles, the bloodstream-specific protein –albeit a mutant version- is able to bind to the cytoskeleton of the procyclic cells.

4.11 Over-expression of CAP5.5 in procyclic cells

In section 4.1, I demonstrated the over-expression of CAP5.5::YFP::Ty in procyclic cells at approximately 1/10th the level of native CAP5.5 using the pDex577 vector. In this section, CAP5.5::YFP::Ty was over-expressed using the pDex777 vector in procyclic parasites in a bid to better understand the effects of excess CAP5.5 in the morphogenesis of the parasite at much higher levels than when pDex577 was used. Transgene expression using the pDex777 vector is mediated by its PFR-2 3' UTR and is higher than when expressed using the pDex577 vector which uses the aldolase 3'UTR (Poon *et al.,* 2012). Over-expression of CAP5.5::YFP::Ty was validated by immunoblotting non-induced and induced (24 hours) populations that had been transfected with the expression construct using monoclonal antibodies against CAP5.5 (Fig. 4.11A). As a comparison of the expression levels, cells over-expressing CAP5.5 using the pDex577 vector under the same experimental conditions are also shown. Quantification of the band intensities revealed that CAP5.5::YFP::Ty was over-expressed at approximately 1.5 times the native CAP5.5 level in the uninduced, when using the pDex777 vector (Fig. 4.11B). Inducible over-expression of CAP5.5::YFP::Ty also led to a 33% decrease in the expression of endogenous CAP5.5 as determined by band intensity (Fig. 4.11B). This might suggest a down regulation of the native CAP5.5 upon over-expression of CAP5.5::YFP::Ty. While over-expression using the aldolase 3'UTR (pDex577 vector) did not exhibit a growth defect, the expression levels achieved by using


A
B
N-P
K-P
length
Induced
Uninduced
C
% cells
K-P (µm)
N-P (µm)
Length (µm)

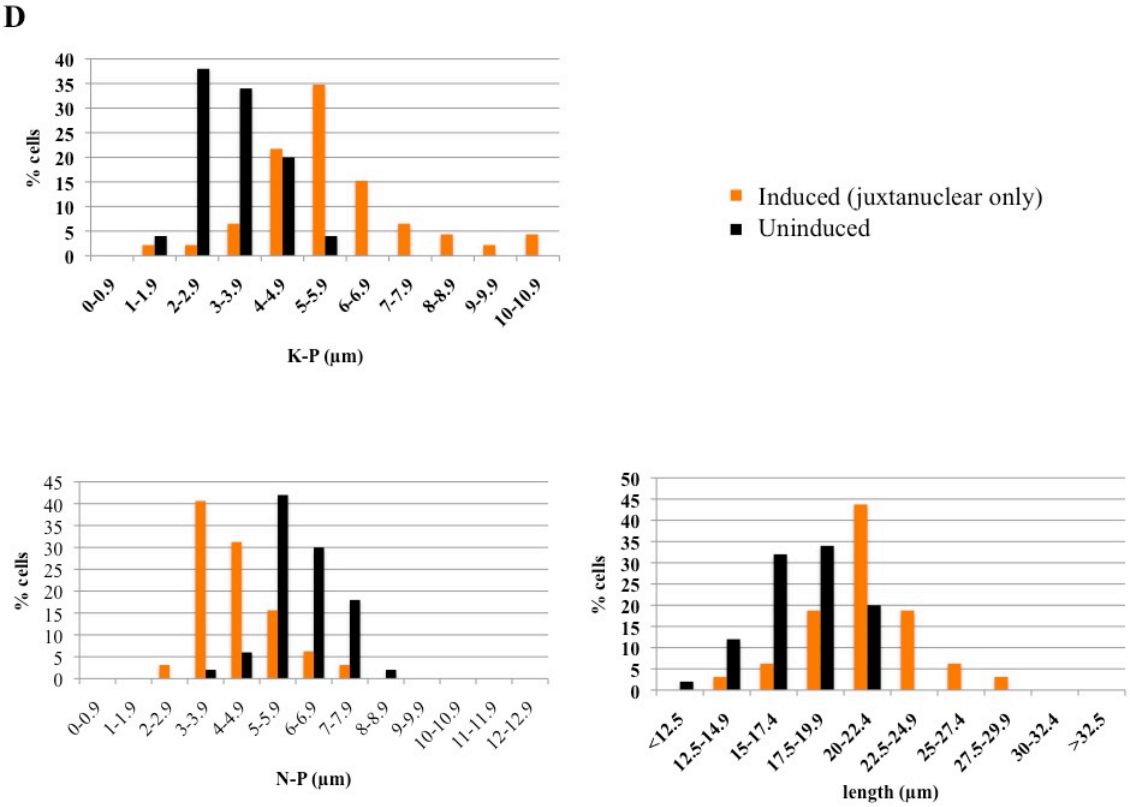
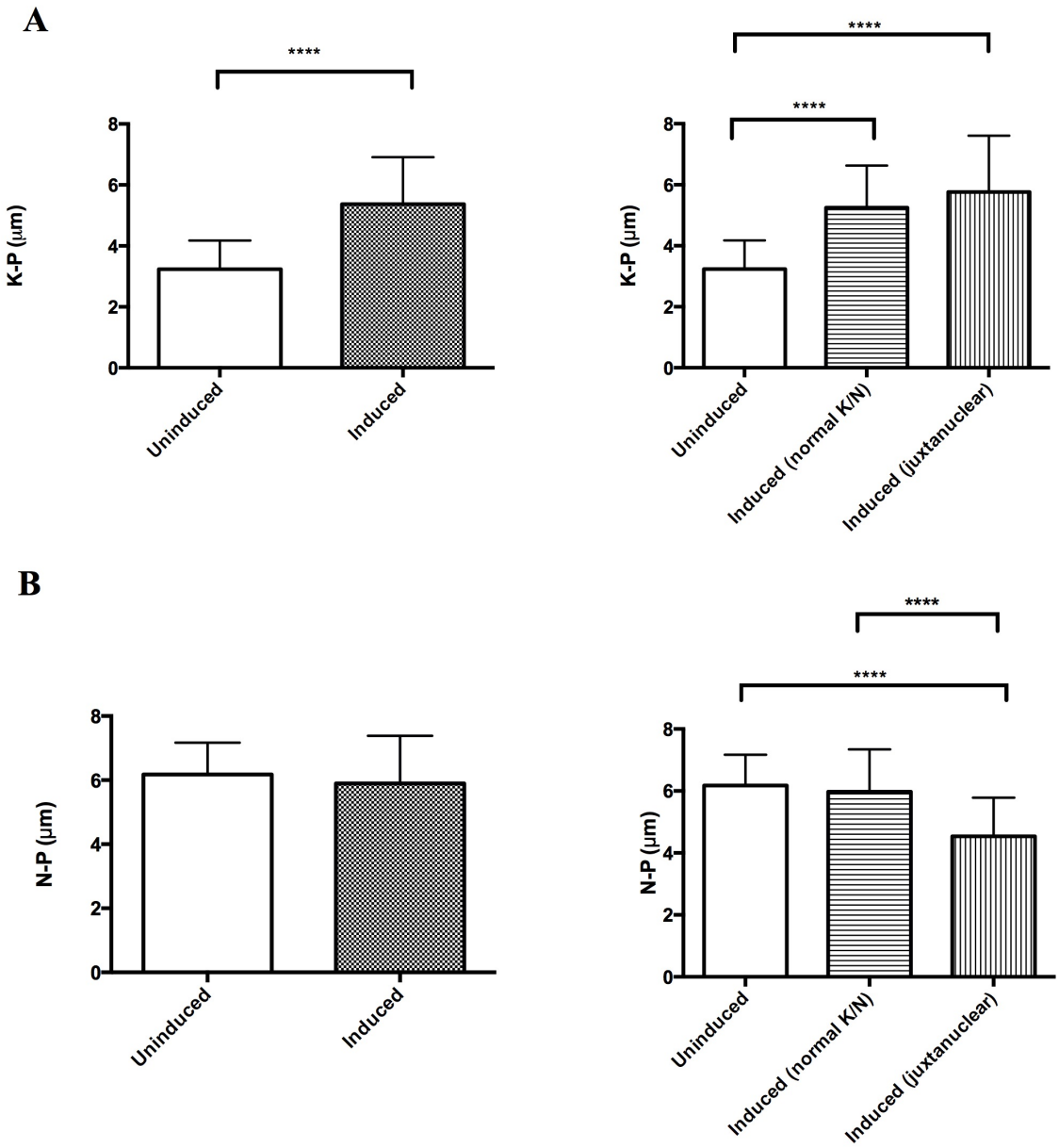


Figure 4.12. Morphometric analyses of procyclic cells over-expressing CAP5.5. A. A phase contrast micrograph of field of view of cells induced for CAP5.5 over-expression for 48 hours. The red arrows highlight the kinetoplast juxta-positioned to the nucleus while the black arrow marks an 'indeterminate' cell. **B**. The criterion for cellular dimensions of 1K1N cells measured are shown. K-P, the distance from the kinetoplast to the posterior end; N-P, distance of the nucleus to the posterior end of the cell; L, cell length including the flagellum. **C**. The frequency distributions of the 1K1N lengths, K-P and N-P distances are plotted for the parental, uninduced cell population (uninduced; n=50) and the induced (48 h) population (induced; n=98). The induced population in C means those with normal and abnormal K/N morphology. **D**. The frequency distributions of the same variables as **C** are plotted of the uninduced and only those induced 1K1N cells with their kinetoplast juxta-positioned to the nucleus (n=48) as shown in A. Scale bar, 3 µm.

the PFR-2 3'UTR (pDex777 vector) showed a noticeable growth defect in induced cells by 48 hours post induction (Fig 4.11C). Over-expression of Ty::YFP::Ty alone in the same vector did not lead to any noticeable growth effect. Cellular localization of whole cells and detergent-extracted cells after 24 hours of induction did not vary from that which was seen in Chapter 4.2, i.e. CAP5.5::YFP::Ty still localized to the cell body – excluded from the flagellum- and such labelling remained after detergent extraction (data not shown).

To investigate the reasons for the growth defect, cells induced for over-expression were analyzed by DAPI staining. This showed an accumulation of abnormal cell types in later stages of the induction (48 and 72 hours post induction) including 1K0N (zoids), 1K2N and cells with multiple kinetoplasts and nuclei (Fig. 4.11D-E). Unfortunately, due to the large number of cells whose kinetoplast/nucleus configurations were not interpretable (called "indeterminate" on graphs), a clearer determination of the cause for the growth defect cannot be made (for an example of an "indeterminate" cell, see Fig. 4.12A). However, one can conclude that CAP5.5's over-expression does lead to a multitude of aberrant cell types with abnormal kinetoplast and nuclear configurations that corroborate the essential nature of CAP5.5 in the morphogenesis of the procyclic form (Hertz-Fowler *et al.,* 2001).

Closer examination of the induced cells with normal kinetoplast and nucleus configurations revealed many with abnormal organelle positioning starting with 24 hours after induction (Fig. 4.12A). At later time points, approximately 31% (n=157/500 at 48 hours post-induction) of the induced cells had such defects. The kinetoplast is normally


A
****
K-P (μm)
8
6
4
2
0
Uninduced
Induced
****
****
K-P (μm)
Uninduced
Induced (normal K/N)
Induced (juxtanuclear)
B
N-P (μm)
Uninduced
Induced
****
****
N-P (μm)
Uninduced
Induced (normal K/N)
Induced (juxtanuclear)

C

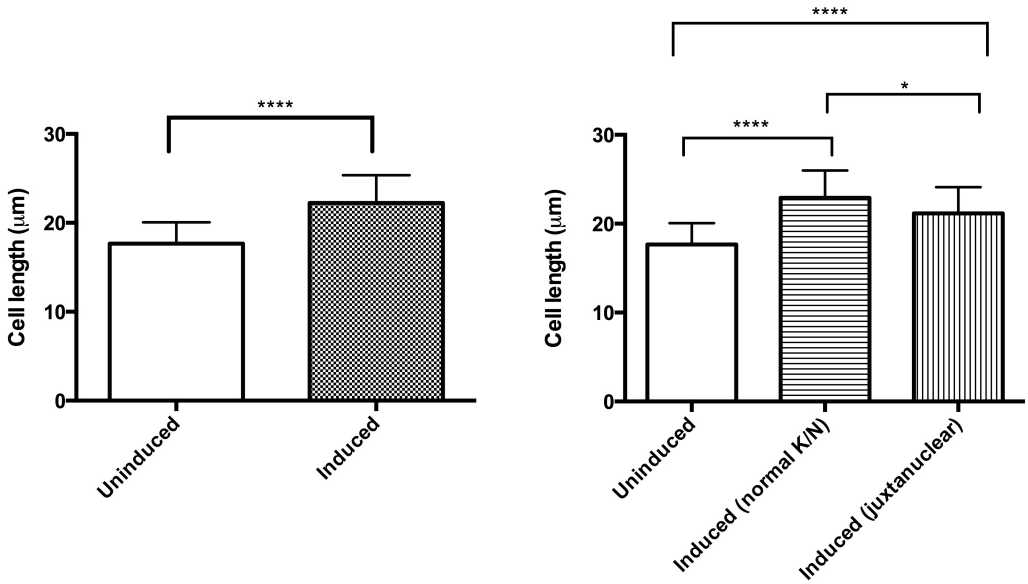


Figure 4.13. Morphometric analyses of 1K1N procyclic cells over-expressing CAP5.5. The mean with its standard deviation for the distance of the kinetoplast to the posterior end (K-P; A), distance of the nucleus to the posterior pole (N-P; B), and the cell lengths (L; C) are shown. The left graphs show uninduced (n=50) and induced cell populations (n=98), where the right graphs recategorize the induced population to those with normal K/N morphology (n=50) and those without (n=48). An unpaired two-tailed t-test was done between uninduced and induced cell populations except for the nucleus to posterior distance (N-P) where a non-parametric Kolmogorov Smirnov test was done (left graphs). An ordinary one-way analysis of variance (ANOVA) and Tukey's multiple comparison test was performed between the three groups on the right graphs. *, p=0.01; ****, p<0.0001.

Table 4.1. Morphometric measurements of 1K1N cells over-expressing CAP5.5 48 hours post induction.

	Length (L) Mean ± SD (µm)	Kinetoplast to posterior (K-P) Mean ± SD (µm)	Nucleus to posterior (N-P) Mean ± SD (µm)
Uninduced	17.66 ± 2.41 µm	3.23 ± 0.94 µm	6.17 ± 0.99 µm
Induced (all)	22.23 ± 2.91 µm	5.37 ± 1.54 µm	5.89 ± 1.49 µm
Induced (normal K/N)	22.92 ± 3.07 µm	5.23 ± 1.35 µm	5.97 ± 1.36 µm
Induced (juxtanuclear)	22.15 ± 2.96 µm	5.58 ± 1.80 µm	4.53 ± 1.24 µm

The table shows the mean values and standard deviations (SD) of cellular dimensions of > 50 1N1K procyclic cells uninduced or induced for over-expression of CAP5.5. The induced cells are further categorized into those with normal K/N morphology and those with kinetoplasts that are juxta-positioned to the nuclei (for images of such cells, please see Fig. 4.12A).

located midway between the nucleus and the posterior end of a 1K1N procyclic cell (Gluenz *et al.,* 2011). However in induced cells, the kinetoplast was juxta-nuclear, with the kinetoplast located to the lateral side or anterior of the nucleus (Fig. 4.12A). The rapid onset of such mispositioning is likely to be a direct consequence of CAP5.5 over-expression. However, the aberrant kinetoplast/nucleus configuration in induced cells may be due to either the mispositioning of the nucleus or the kinetoplast. In order to resolve this issue, DAPI staining of 1K1N cells were analyzed. 1K1N cells were chosen as at any given time, such cells form the majority of an unsynchronized culture and also to facilitate morphometric analyses. The distance from the kinetoplast and nucleus to the posterior pole of 1K1N uninduced cells and cells that had been induced for 48 hours, regardless of kinetoplast/nucleus positioning, were measured and analyzed (Fig. 4.12B-C; n $\geq$50). All of these parameters showed a normal distribution in both induced and uninduced cells except for the induced (all) N-P distance, which showed a possible bimodal distribution, which might account for the two distinct populations observed: cells with and without their kinetoplasts juxta-postioned to the nucleus (as seen in Fig. 4.12A). The mean values with standard deviation of the distance from the kinetoplast to the posterior end (K-P), nucleus to the posterior end (N-P), and cell length (L; including the flagellum) of 1K1N populations that had been induced for 48 hours are tabulated in Table 1. Of the 1K1N induced population, as not all 1K1N cells had organelle mispositioning, the induced population was further categorized into those with normal and abnormal N/K positioning, where the juxta-nuclear kinetoplast classifies as abnormal (Fig. 4.12D).

The kinetoplast to posterior end distance (K-P) increased by approximately 2.1 μm ($p<0.0001$; two-tailed t-test) in induced cells, regardless of the kinetoplast/nucleus

configuration (Fig. 4.13A; Table 1). The increase in distance of the kinetopast from the posterior end may be due to the active microtubule extension during cell elongation that takes place at the posterior ends of the cell and not necessarily due to the anterior movement of the kinetoplast (Sherwin & Gull, 1989b; Robinson *et al.,* 1995).

Although, there was no significant difference between the nuclear distance to the posterior pole of the cell (N-P) between the induced and uninduced cell population ($p=0.24$; Kolmogorov Smirnov test), there was a decrease in N-P distance of approximately 36% (1.64 μm; $p<0.0001$; Tukey test) in those induced cells with abnormal K/N configuration relative to uninduced cells (Fig. 4.13B). Among the induced cells, the N-P distance was significantly different between those with different K/N morphologies (Fig. 4.13B; $p <0.0001$; two tailed t-test). On average, cells from the total induced 1K1N populations were 25% longer (~4.57 μm) than the uninduced 1K1N cell population ($p <0.0001$; two-tailed t-test). Comparison of the uninduced 1K1N cells with induced cells exhibiting normal or abnormal K/N morphology revealed an approximately similar difference (Fig. 4.13C; Table 1). Taken together, this suggests that the posterior ends of cells, upon CAP5.5 over-expression, are increasing in length, probably due to the active extension of posterior ends. Microtubule extension at the posterior end should normally increase the N-P distance (and K-P distance), however, a significant decrease in the distance of the nucleus from the posterior pole of the induced cells was noticed. Therefore, this is highly suggestive of a nucleus mis-positioning towards the posterior end of the parasite, whereby it appears juxta-positoned to the normally posterior kinetoplast. To date, little is known about how the position of the nucleus is determined in the cell division cycle and linkages between the sub-pellicular cytoskeleton and the

nucleus have not yet been observed in structural studies. RNAi against *CAP5.5* and *CAP5.5V* have also observed nuclear mis-positioning at later points in the cell cycle, so, might CAP5.5 and/or the binding partners of other cytoskeleton associated proteins, be involved in mediating the securing of certain organelles to the cytoskeleton?

4.12 Summary

In this chapter, a series of CAP5.5 deletion constructs using the pDex577 vector were expressed in procyclic *T. brucei* to determine the molecular determinants of its trafficking to the sub-pellicular cytoskeleton (summarized in Fig. 4.14). While, N6::YFP::Ty localized to the cytosol and the flagellum, N37::YFP::Ty, DI::YFP::Ty , DI-III::YFP::Ty and CAP5.5::YFP::Ty all localized to the pellicular membrane (with labelling excluded from the flagellum). Cells over-expressing DI-II::YFP::Ty exhibited non-uniform cytosolic localization. Taken together, the data suggested that within the stretch of the N-terminal 37 amino acids lie molecular determinants that possibly retain CAP5.5 in the cell body, as N6::YFP::Ty localize predominantly to the flagellum. Additionally, labelling of all the aforementioned fusion proteins except for CAP5.5::YFP::Ty were susceptible to non-ionic detergent extraction, suggesting that the domain mediating CAP5.5's association with the sub-pellicular cytoskeleton may be domain IV. Temporal analysis of CAP5.5 over-expression suggested that CAP5.5 and presumably other MAPs may be added from the more dynamic, posterior pole of the procyclic cell, especially those in earlier stages of the cell division cycle. Ectopic expression of N-terminally GFP-tagged CAP5.5 and CAP5.5V using the pDex777 vector, where the acylation is hypothesized to be disrupted, caused growth defects in both

CAP5.5	3' UTR	Localization
FL::YFP::Ty	PFR-2; aldolase	Subpellicular cyto.+ pellicular membrane
GFP::FL	PFR-2	Subpellicular cyto. + pellicular membrane
D I-III::YFP::Ty	Aldolase	Pellicular membrane*
D I-II::YFP::Ty	Aldolase	Cell body/flagellum*
D I::YFP::Ty	Aldolase	Cell body + pellicular membrane
N37::YFP::Ty	Aldolase	Cell body + pellicular membrane
N6::YFP::Ty	Aldolase	Cell body + flagellum
CAP5.5V		
GFP:: FL	PFR-2	Subpellicular cyto. + pellicular membrane

Figure 4.14. A summary of the localization of the ectopic expression constructs used in this study. The localization profiles here refer to those seen in procyclic cells upon tetracycline induction for 24 hours. * refers to heterogeneity in the localization profiles observed.

procyclic and bloodstream forms. Additionally, clear labelling of CAP5.5V fusion protein was seen in the cytoskeletons of procyclic cells, suggesting that the bloodstream-specific CAP5.5V can associate with the cytoskeleton of a different developmental stage parasite. Over-expression of CAP5.5::YFP::Ty using the pDex777 vector led to a noticeable growth phenotype, that was characterized by the accumulation of abnormal cell types, abnormal organelle positioning, most prominently the juxta-nuclear positioning of the kinetoplast and statistically significant elongation at the posterior ends of the induced cells.

5. Discussion

5.1 Introduction

Calpains - calcium-regulated cysteine proteases - can be broadly categorized into typical and atypical calpains, with the distinction being the presence of a calcium-binding penta-EF hand (PEF) domain in typical calpains. Typical calpains are largely restricted to metazoans, with mammals encoding 15 calpain genes and *C. elegans*, 12 (reviewed in Goll *et al.,* 2003). Outside the metazoa, all sequenced organisms such as *Plasmodium falciparum, Cryptosporidium parvum*, and *Entamoeba histolytica,* encode only one calpain-like gene except for kinetoplastid parasites (Gardner *et al.,* 2005; Anderson & Loftus, 2005; Abrahamsen *et al.,* 2004; Croall & Ersfeld, 2007).

Systematic analyses of the genomes of three kinetoplastid parasites revealed the presence of a large number of calpain-like proteins: 18 in *T. brucei,* 24 in *T. cruzi* and 27 in *L. major* (Ersfeld *et al.,* 2005). Others estimate an even larger expansion of calpain-encoding genes in these parasites - almost double the published estimates (Olego-Fernandez, Oxford DPhil thesis 2010). Nonetheless, to my knowledge, this represents the largest number of calpains (typical or atypical) in any sequenced organism containing calpain-encoding genes and therefore, may indicate trypanosomatid-specific functions of this gene family. Discovery of such expanded calpain gene families not only in lower eukaryotes but also in animals, fungi, plants and bacteria demands a revision of the terms "typical" and "atypical" used to conventionally describe calpains, when these "atypical" calpains appear to be the norm and not necessarily the exception (Zhao *et al.,* 2012; Campbell & Davies, 2012). This thesis shall maintain the use of the standard terminology in order to avoid confusion. However, a more befitting term may be classical calpains in

reference to typical calpains as these calpains are similar to CAPN1 and CAPN2, the first calpains to be characterized (Campbell & Davies, 2012).

Since all kinetoplastid calpains are atypical, PEF domain acquisition in typical calpains must have occurred fairly recently by gene fusion during animal evolution (Ohno *et al.,* 1984). The expansion of this gene family in kinetoplastid parasites is hypothesized to have occurred as a result of a multiple series of gene duplication, which may have facilitated the evolution of novel functions (Ersfeld *et al.,* 2005; Ohno, 1970). Detailed perspectives on the evolution of calpains in lower eukaryotes are presented in these papers (Jékely & Friedrich, 1999; Zhao *et al.,* 2012).

Typical calpains have been implicated in a range of physiological processes that include cytoskeletal remodelling, signal transduction, cell motility, cell cycle progression, and differentiation (Lebart *et al.,* 2006; Goll *et al.,* 2003). Whereas some of the roles of atypical calpains are myofilament turnover, neurodegeneration, cytoskeletal remodelling, sex determination, and adaptive responses to alkaline environments (Ono & Sorimachi, 2012; Olego-Fernandez *et al.,* 2009).

Precise functions of kinetoplastid calpains are not yet known. Proteomic analyses of benznidazole-resistant *T. cruzi* epimastigotes showed SKCRPx.1 (an ortholog of *L. major* SMP-1) as being over-expressed (Andrade *et al.,* 2008). Application of various types of stress (nutritional, temperature, pH) to *T. cruzi* epimastigotes found TcCALPx11 as being upregulated (Giese *et al.,* 2008). In *T. brucei,* knockdown of *CALP1.2* and *SKCRP1.4* in bloodstream and procyclic cells respectively were lethal (Subramaniam *et al.,* 2006; Liu *et al.,* 2010). However, RNAi of other calpains found on *T. brucei*

chromosome 1 such as *CALP1.1*, *CALP1.3* and *CALP1.4* in bloodstream cells did not exhibit a growth phenotype in a systematic RNAi screen of chromosome 1 genes (Subramaniam *et al.,* 2006). RNAi-based characterization of CAP5.5 and CAP5.5V revealed severe perturbations of cytokinesis accompanied by organelle mis-positioning and disordered cytoskeleton organization in procyclic and bloodstream cells (Olego-Fernandez *et al.,* 2009).

In this thesis, one of the issues I investigated was whether the physiological functions of kinetoplastid calpains, as represented by CAP5.5 and CAP5.5V, were dependent upon an intrinsic proteolytic activity (Chapter 3). I also explored the molecular determinants of CAP5.5's localization to the sub-pellicular cytoskeleton and dominant-negative analyses of CAP5.5 over-expression (Chapter 4). Here, I will build upon the discussion I initiated in the respective chapters and review the results in the context of the current literature on calpains and trypanosome cell biology.

5.2 Aggregation of recombinant calpains

In the first sections of Chapter 3, I discussed the expression and purification of recombinant calpains, primarily the role of aggregation in impeding structural and biochemical characterization of CAP5.5 and CAP5.5V. Aggregation of soluble full-length CAP5.5 and CAP5.5V and the inability to reverse this process despite further optimization led me to pursue the expression of CAP5.5V's individual domains or combinations thereof. Following the determination of an aggregation mediated by disulfide bond and hydrophobic association, I was able to reverse it by reducing the ionic strength of buffers to generate monodisperse and purified CAP5.5V domain II for

biochemical characterization. However, extrapolation of these *in vitro* results -especially the aggregation- to trypanosomal cell biology is naturally speculative, as large molecular-weight CAP5.5/CAP5.5V aggregates are artifacts of purification promoted by the chemical reactions of calpains in artificial conditions.

Aggregation of calpains has been frequently noted and impeded structural characterization of typical calpains for many years (Yoshizawa *et al.*, 1995; Elce *et al.*, 1997). In the case of the dimeric rat m-calpain, Ca^{2+} activation led to subunit disassociation, exposing the hydrophobic dimerization interface causing oligomerization (Pal *et al.*, 2001). The structures of mammalian calpains were solved using ethylenediaminetetraacetic acid (EDTA) to chelate calcium or endogenous inhibitors such as calpastatin (Moldoveanu *et al.,* 2002; Hanna *et al.*, 2008). While the aggregation observed for typical calpains is mediated by calcium activation, we know that CAP5.5/CAP5.5V lack three out of the five conserved acidic residues required for calcium binding (Fig. 3.2). Nonetheless, the purification protocols used here employed the use of EDTA as chelators to no avail, suggesting an altogether different cause of aggregation for the full-length CAP5.5 and CAP5.5V.

Aggregation is inherently an equilibrium process and in the absence of autolysis, should be fully reversible (Pal *et al*., 2001). Analyses of CAP5.5V domain II aggregates in buffers of different ionic strengths revealed that aggregation was reversible in buffers with low physiological salt concentrations as determined by light scattering (Chapter 3). Additionally, the inhibition and reversal of aggregation by lower salt concentration strongly suggested that hydrophobic interactions were involved. This observation is based on the rationale that decreasing the ionic strength promotes an increase in water

solvation layers around a protein molecule thereby mitigating hydrophobic-mediated aggregation. However, this approach did not work with the purification of the full-length calpains suggesting a more complicated and multifactorial aggregation for the full-length calpains.

5.3 CAP5.5V D-II has no detectable protease activity

The determination of whether CAP5.5 and CAP5.5V have protease activity is important not only to our appreciation of kinetoplastid calpains in general but also for a better understanding of the essential roles CAP5.5/CAP5.5V play in the proper morphogenesis of *T. brucei* (Olego-Fernandez *et al.,* 2009). Recombinant protease cores of typical calpains such as rat CAPN1 have been demonstrated to retain proteolytic activity characteristic of the full-length calpains, even when expressed as individual units (Moldoveanu *et al.,* 2002; Moldoveanu *et al.,* 2004a). Here, I showed that the domain II of CAP5.5V, which corresponds to the protease core domain of typical calpains, does not have any detectable autolytic activity or activity against universal protease substrates such as casein *in vitro* (Chapter 3.5-3.7). Equal concentrations (10 μg) of substrate and recombinant calpains were used. Therefore, the lack of any detectable proteolytic activity observed for CAP5.5V domain II suggest by extension proteolytic inactivity for the CAP5.5V and CAP5.5 (on the basis of their homology) full-length proteins (Fig. 2.1).

Given the highly conserved nature of catalytic residues, it may be difficult to envisage a classical cysteine protease-type cleavage mechanism with the putative catalytic triad configurations (Ser_{291}-Tyr_{454}-Asn_{476}) of CAP5.5V. However, other possibilities existed. Firstly, it is possible that CAP5.5V functions with a different

nucleophile (S291 instead of the canonical cysteine) as was observed with poliovirus-type picornain C3 peptidases. These proteases belong to the serine protease superfamily but operate with a cysteine nucleophile in cleaving viral proteins instead of the characteristic serine (Sarkany & Polgar, 2003). Secondly, the tyrosine (Y454) in the catalytic triad itself can act as an alternate catalytic nucleophile, much like in short-chain dehydrogenases/reductases, which operate with a similar catalytic tetrad comprising of Asn-Ser-Tyr-Lys (Kavanagh *et al.,* 2008). Therefore, I had sought to establish the proteolytic profile of these calpains by analysis of recombinant CAP5.5V domain II.

Even if CAP5.5V and CAP5.5 may not be bona fide proteases, as my results suggest, one can envisage a range of possible roles based on observations of other pseudoenzymes and non-enzymatic homologues. Pseudoenzymes and non-enzymatic homologues –proteins with major substitutions in their catalytic sites - can be found in the proteomes of both prokaryotes and metazoans, where they have evolved to acquire novel functions in regulatory networks that do not necessarily rely on proteolytic processing (Todd *et al.,* 2001). A large-scale analysis of non-enzymatic homologues, not unlike CAP5.5/CAP5.5V, found that such proteins are very often found in signalling pathways, and hardly at all in nuclear processes (Pils & Schultz, 2004). A mechanistic explanation could be that alterations in the functions of nuclear enzymes would be fatal whereas signalling molecules may be able to tolerate greater mutations, thereby increasing the complexity of regulatory networks (Pils & Schultz, 2004).

Indeed, regulation of active proteases has been shown to occur through related non-enzymatic homologues either by directly modulating the activity of the active homologue or by indirectly countering the effects of the active enzyme such as through

substrate sequestration and subsequent substrate degradation (Pils & Schultz, 2004; Todd *et al.,* 2001). For example, cellular FLICE-inhibitory protein (c-$FLIP_L$) in the extrinsic cell death pathway activates its active homologue, procaspase-8 through heterodimerization of their respective protease domains (Yu *et al.,* 2009). On the other hand, in the epidermal growth factor (EGF) pathways of *Drosophila* and mammalian cells, rhomboid pseudoproteases inhibit rhomboid-dependent signalling by promoting the sequestration and subsequent ER-associated degradation of potential rhomboid substrates (Zettl *et al.,* 2011). An explanation for the latter mechanism could be that non-enzymatic homologues may still be able to bind to substrates but not be able to cleave them suggesting a shift from substrate catalysis to substrate binding in non-enzymatic homologues (Lamb *et al.,* 2000).

In mammals, CAPN6, like CAP5.5 and CAP5.5V, is an atypical calpain with major substitutions in its catalytic triad residues (Tonami *et al.,* 2007). Characterization of CAPN6 found it to be a microtubule-stabilizing protein, with CAPN6 over-expression causing the formation of microtubule bundles and RNAi against *CAPN6* causing membrane ruffling, actin reorganization and microtubule instability (Tonami *et al.,* 2007). Given that the typical substrates of mammalian calpains are cytoskeletal proteins, it has been suggested that loss of protease activity may have been a necessary step in the evolution of CAPN6's function as a microtubule-stabilizing protein involved in the regulation of cytoskeletal organization (Goll *et al.,* 2003; Olego-Fernandez *et al.,* 2009). Likewise for CAP5.5 and CAP5.5V, the loss of proteolytic activity may have prompted an evolution of the functions of these cytoskeleton-associated proteins. Their localization profiles and association with the plasma membrane are highly conducive to a regulatory

role given that cytoskeleton can also act as a conduit for signal transduction via linkages to the membrane and membrane receptors (Barth *et al.,* 1997; Schoenwaelder *et al.,* 1999). Additionally, attachment of lipophilic moieties to proteins facilitates protein interactions, membrane binding and targeting, and intracellular signalling (Resh, 1996).

5.4 Localization and targeting of CAP5.5

CAP5.5 belong to a group of dually acylated proteins that do not localize to the flagellum, the default localization of such proteins (Emmer *et al.,* 2009). CAP5.5 and CAP5.5V have been demonstrated to associate with the sub-pellicular cytoskeleton and not the axonemal microtubules of the flagellum (Hertz-Fowler *et al.,* 2001; Olego-Fernandez *et al.,* 2009). Other dually acylated proteins in that category include hydrophilic acylated surface protein B of *L. major,* which is an extracellular protein, and phosphoinositide-specific phospholipase C in *T. cruzi,* which localize to the plasma membrane in amastigotes and to the cytosol in epimastigotes (Denny *et al.,* 2000; Furuya *et al.,* 2000; Okura *et al.,* 2005). Therefore, given their localizations, there are likely to be additional signals modulating the localization of these proteins. To investigate what elements in CAP5.5's sequence are involved in targeting it to the sub-pellicular cytoskeleton, a series of truncation mutants were generated. In Chapter 4, I discussed the subcellular localization and biochemical characteristics of several such CAP5.5 mutants.

Analysis of CAP5.5's N-terminus revealed that N6::YFP::Ty, comprising the 6 amino-acid long dual acylation motif, localized predominantly to the flagellar matrix in procyclic cells whereas N37::YFP::Ty and domain I::YFP::Ty localized to the cell body and was absent from the flagellum (Fig. 4.4 - 4.6). N37::YFP::Ty and DI::YFP::Ty also

showed a clear accumulation around the periphery of the cell, suggestive of an association with the pellicular membrane. This clear difference in localization followed the addition of only 29 amino acids to N6::YFP::Ty. One explanation might be that this stretch of amino acids (in N37::YFP::Ty) contributes to a different interaction with the lipid environment of the flagellum as compared to N6::YFP::Ty, as N37::YFP::Ty and domain I::YFP::Ty are retained in association with the pellicular membrane with its translocation to the flagellum prevented.

In kinetoplastid parasites, proteins modified with lipophilic moieties are predominantly targeted to the flagellar membrane, where there is an enrichment of sterols, glycolipids, sphingolipids, all of which are components of canonical lipid rafts in contrast to the microenvironment of the cell body (pellicular) membrane (Chailley, 1985; Tetley, 1986; Kaneshiro *et al.,* 1984). The third domain of the kinetoplastid cell surface is the flagellar pocket membrane. Disruption of lipid rafts by cholestrol depletion or by fluidization of the membrane abolishes the native flagellar membrane localization of dually acylated proteins like the flagellar calcium-biding protein (FCaBP) in *T. cruzi* (Tyler *et al.,* 2009). Therefore, association of dually acylated proteins with such lipid rafts has been suggested as a possible mechanism for flagellar membrane trafficking in kinetoplastids (Fridberg *et al.,* 2006; Tyler *et al*., 2009). Other well studied dually acylated flagellar proteins include calflagins in *T. brucei* and SMP-1 in *L. major* (Emmer *et al.,* 2009; Tull *et al.,* 2004).

However, while dual acylation is necessary, it is certainly not sufficient for any membrane localization. Additional stabilizing interactions, be they electrostatic or hydrophobic, are required (Hurley & Misra, 2000). In FCaBP, conserved lysines in the

first 13-24 amino acids were shown to confer localization specificity and lipid raft affinity (Maric *et al.,* 2011). Similarly, the flagellar localization of dually acylated *Tb*CALP1.3 has been demonstrated to require a 35 amino acid long C-terminal stretch of basic residues in addition to the dual acylation (Liu *et al.,* 2010). On the other hand, the flagellar membrane localization of small myristoylated protein -1 (SMP-1) in *L. major* has been demonstrated to be largely mediated by hydrophobic interactions in addition to the dual acylation (Tull *et al.,* 2004).

CAP5.5 and CAP5.5V contain two possible motifs and domains with sequence homology to domains known to stabilize membrane interactions. Firstly, that N-terminal 37 amino acids stretch (N37::YFP::Ty) contains a proline-rich region immediately following the myristoylation/palmitoylation signal. The sequence PxxP conforms to the binding sequence motif found in many ligands of SH3- type domains (Ren *et al.,* 1993; Yu *et al.,* 1994). Proline-rich regions are recognized as mediating protein-protein interactions, particularly at the plasma membrane interface (Buday *et al.,* 1999; Niebuhr *et al.,* 1997). Therefore, the retention of N37::YFP::Ty and not N6::YFP::Ty in the pellicular membrane could be due to interactions between the polyproline stretch and the pellicular membrane. Secondly, CAP5.5 and CAP5.5V domain III bear a ~19% sequence homology to the C2-like domain of typical calpains. C2L domains are found in a number of enzymes such as protein kinase C and phospholipases that transiently bind to membranes (Corbalan-Garcia *et al.,* 2010; Rizo *et al.,* 1998). Therefore, this domain could further stabilize CAP5.5/CAP5.5V membrane association in a non-specific hydrophobic interaction. In accordance with such hypothesis, CAP5.5 domain I-

III::YFP::Ty did show a clear accumulation in the cell's periphery, highly suggestive of pellicular membrane association.

I must note here that domain I-II::YFP::Ty localized to the cytoplasm without any clear peripheral accumulation and localized to a lesser extent in the flagellum as well. However, only a small minority of transfected cells were YFP positive, which I hypothesized to indicate that the fusion protein may be misfolded and not processed properly, thereby targeted for degradation by the cellular machinery.

Additionally, procyclic cells expressing YFP fusion proteins (N6, N37, domain I, domain I-II, domain I-III) were susceptible to *in situ* extraction with a non-ionic detergent suggesting that domain IV may be the domain mediating cytoskeleton association as CAP5.5 has been shown to bind to the cytoskeleton by both biochemical and structural analyses (Hertz-Fowler *et al*., 2001). Such an observation is especially interesting when considering that the major difference between the paralogs CAP5.5 and CAP5.5V is domain IV. As these two cytoskeleton-associated proteins perform analogous roles in different lifecycle stages, it may be tempting to assume that the difference in domain IV accommodates the differences in post-translational modifications and composition of the cytoskeleton of the two life forms (Benz *et al.,* 2012). However, over-expression of CAP5.5V, albeit with a N-terminal GFP tag, was able to associate with the sub-pellicular cytoskeleton of procyclic cells suggesting that such an assumption is not true (Fig. 4.10).

In higher eukaryotes, several examples of membrane proteins being retained in specific membrane subdomains that require cytoskeletal tethering are known such as the erythrocyte anion exchanger band 3, neuronal glycine receptors, and Na+, K+-ATPases

(Bennett & Stenbuck, 1979; Kirsch *et al.,* 1991; Morrow *et al*., 1989). In trypansomatid protozoans, there are examples emerging that are compatible with a model of membrane proteins being targeted to the cell body that also require an association with the sub-pellicular cytoskeleton.

In *L. enriettii*, the major glucose transporter exists in two isoforms: iso-1, which localizes to the flagellar membrane and does not associate with the cytoskeleton and iso-2, which localizes to the pellicular membrane and associates with the underlying cytoskeleton (Snapp & Landfear, 1997). They differ only in their N-terminal domains. Deletion of 25 or more amino acids from the N-terminus retargets iso-1 from the flagellum to the pellicular membrane, where the mutant iso-1 protein associates with the cytoskeleton. This suggested that cytoskeletal association acted as a tether to facililate the pellicular membrane localization of iso-2 but also that the flagellar isoform, iso-1, is prevented from establishing such an association due to the physicial sequestration into a compartment devoid of sub-pellicular microtubules (Snapp & Landfear, 1997; Nasser & Landfear, 2004).

Application of such a trafficking model to CAP5.5 or CAP5.5V would suggest that the association of these calpains to the sub-pellicular cytoskeleton serves as an anchor to facilitate their pellicular membrane interactions. However, the temporal order of CAP5.5's association with cytoskeleton and the membrane remains to be investigated.

5.5 Posterior-end incorporation of CAP5.5 into the procyclic cytoskeleton

CAP5.5 over-expression under the control of an aldolase 3'UTR did not cause any detectable growth defect. Additionally, CAP5.5::YFP::Ty localized to the sub-pellicular cytoskeleton like native CAP5.5. Given the absence of a dominant negative effect, the data suggested the likelihood of this fusion protein being a good model for its wild-type counterpart. Therefore, I further analyzed cells over-expressing CAP5.5::YFP::Ty in order to better understand CAP5.5 expression.

In Chapter 4.3, I showed how CAP5.5::YFP::Ty labeling of the sub-pellicular cytoskeleton was preferentially restricted to the posterior ends of cells at the earlier stages of induction and how such a labeling became more widespread with time until the entire sub-pellicular cytoskeleton was labelled. Additionally, this posterior region was YL1/2 positive, which recognizes dynamic microtubules (Kilmartin *et al.,* 1992).

Trypanosomal microtubules can be found in two major organizations. The microtubules in the sub-pellicular corset are oriented such that the fast-growing, dynamic and plus ends of the microtubules are located at the posterior end, where both initiation of cell elongation and intercalation of new microtubules take place during the early stages of the cell cycle (Sherwin, & Gull, 1989b; Robinson *et al.,* 1995). This is in contrast to the axonemal microtubules in the flagellum, where the polarity is reversed as these microtubules have their plus ends at the anterior end of the cell or the distal end of the flagellum. The flagellar attachment zone (FAZ) is a complex structure located in a gap between sub-pellicular microtubules that comprise of the FAZ filament and the microtubule quartet. This quartet, with its nucleation point being close to the basal

bodies, have an opposite polarity to the sub-pellicular microtubules (Sherwin & Gull, 1989; Robinson *et al.,* 1995).

The posterior-specific CAP5.5::YFP::Ty localization observed complements the initial RNAi phenotypes of CAP5.5 and other microtubule-associated proteins such as auxin-induced in root cultures 9 (*Tb*AIR9) and whole cell body (WCB) proteins in *T. brucei* (Olego-Fernandez *et al.,* 2009; Baines & Gull, 2008; May *et al.,* 2012). This RNAi phenotype entailed the preferential depletion of these proteins from the posterior ends of the procyclic cell, where the microtubules are most dynamic. Taken together this implies that CAP5.5, and presumably other microtubule-associated proteins, are added to the newly re-modelled sub-pellicular microtubule array during cell elongation at its dynamic posterior ends. Some MAPs like CAP5.5, then spread from the posterior to the anterior ends of the cell, having been built in from the cell's posterior ends to eventually associate with the entire sub-pellicular cytoskeleton, as is observed in the later stages of the CAP5.5::YFP::Ty over-expression.

This is in contrast to the proposed models of flagellar protein transport and flagellar assembly. While the flagellum is nucleated from the basal body, flagellar proteins need to be synthesized in the cyotoplasm as the flagellar matrix lacks ribosomes (Rosenbaum *et al.,* 1999). A complex system called intraflagellar transport (IFT) is responsible for transporting cargo from either the base to the tip of the flagellum (anterograde transport) or vice versa (retrograde transport) and is facilitated by kinesins and dyneins (Cole *et al.,* 1998; Kozminski *et al.,* 1993; Pazour *et al*., 1999). Flagellar proteins are transported to the flagellum base driven by localization signals and motifs among other possible mechanisms (Bastin *et al.,* 1999; Ersfeld *et al.,* 2001). Newly

synthesized flagellar proteins must be transported to the polymerization-dynamic distal end before being incorporated into the major flagellar structures such as the axoneme and the paraflagellar rod (Bastin *et al.,* 1999; Kohl *et al.,* 2005).

5.6 Cell cycle-regulated incorporation of CAP5.5 into the procyclic cytoskeleton

I was also interested in understanding whether CAP5.5's incorporation into the sub-pellicular cytoskeleton was cell cycle regulated. At early stages of induction, cells with posterior-end restricted CAP5.5::YFP::Ty expression were mostly at an earlier stage [1K1N; 1K1N (elongated K)] of the cell cycle with a noticeable absence of 2K2N cells that were YFP-positive, a proxy for CAP5.5::YFP::Ty incorporation into the cytoskeleton. 2K2N cells have completed mitosis and are poised for cytokinesis (Woodward & Gull, 1990). At later time points, with the progression of cells through the cell division cycle, CAP5.5::YFP::Ty + cells at later stages of the cycle (2K1N, 2K2N) increased with a corresponding decrease in cells with 1K1N and 1K1N (elongated K) configurations.

Analysis of the current data suggests a possible cell cycle regulation of CAP5.5 incorporation into the cytoskeleton, with CAP5.5 being added to the sub-pellicular cytoskeleton from the posterior end of the cell at earlier stages of the cell cycle. This suggests an elegant mechanism for cell cycle regulated incorporation of MAPs like CAP5.5 into the cytoskeleton.

5.7 Functional characterization of CAP5.5

RNAi against *CAP5.5* and *CAP5.5V* in procyclic and bloodstream cells respectively have demonstrated the two calpains to be essential for the proper morphogenesis of *T. brucei* (Olego-Fernandez *et al.,* 2009). In both life-cycle stages, RNAi led to a defective growth phenotype accompanied by the accumulation of cells with abnormal kinetoplast/nucleus configurations such as anucleate "zoids" (1K0N) and binucleate 1K2N cells. This is consistent with organelle mis-positioning or aberrant cleavage furrow ingression during cytokinesis. Additionally, microtubule bundling beneath the pellicular membrane, particularly in the posterior end of the cell and aberrant inter-microtubule spacing was observed, suggesting perhaps a structural role in maintaining these inter-microtubule linkages.

In a corollary experiment described in chapter 4, I discussed the inducible over-expression of CAP5.5::YFP::Ty in procyclic cells regulated by the PFR-2 3'UTR, where I also observed a retarded population growth rate and the production of a multitude of aberrant cell types with abnormal kinetoplast and nuclear configurations that corroborate the essential nature of CAP5.5 in the morphogenesis of the procyclic form (Fig. 4.12-4.13; Hertz-Fowler *et al.,* 2001). A clearer determination of the cause for the growth defect could not be made owing to the large number of morphologically aberrant cells with indeterminable kinetoplast/nucleus configurations.

In procyclic trypanosomes, over-expression of CAP5.5::YFP::Ty significantly affected the positioning of the nucleus, resulting in its relocation to the posterior of the cell juxta-positioned to the kinetoplast, a positioning reminiscent of the epimastigote life cycle stages in the tsetse fly (Sharma *et al.,* 1998). Interestingly, such a phenotype

mimics the RNAi phenotype of another microtubule-associated protein, *Tb*AIR9 in procyclic *T. brucei* (May *et al.,* 2012). However, depletion of TbAIR9 in bloodstream cells had little effect on nuclear positioning but resulted in aberrant cytokinesis (May *et al.,* 2012).

Additionally, the cell lengths of CAP5.5::YFP::Ty induced cells were profoundly longer than uninduced cells probably due to active microtubule extension at the posterior ends of the cells. One would expect that microtubule extension at the posterior end would increase the nucleus to posterior end (N-P) distance and kinetoplast to posterior distance (K-P), however, a significant decrease in the distance of the nucleus from the posterior pole of the induced cells was noticed. Therefore, this is highly suggestive of a nucleus mis-positioning towards the posterior end of the parasite, whereby it appears juxta-positoned to the normally posterior kinetoplast.

Although the morphometric analyses were limited to 1K1N cells, which form the majority of an unsynchronized culture at any given time, cells later in the cell cycle (such as 2K1N, 2K2N) also exhibited similar nuclear mis-positioning defects (data not shown). The nuclear mis-positioning defect was observed in cells as early as after 24 hours of induction. While, I hypothesize that such an early phenotype may correlate to a direct consequence of CAP5.5 over-expression, whether nuclear mis-positioning leads to the aberrant cell phenotypes observed or vice versa remains unknown.

Trypanosomes exhibit a highly ordered spatial and temporal organelle positioning and segregation that have been exploited in classifying the different life cycle stages and cell cycle stages of parasites (Vickerman, 1976; Woodward & Gull, 1990). For instance,

the kinetoplast, which is positioned midway between the nucleus and the posterior end of procyclic cells, adopt an anterior position relative to the nucleus (due to nuclear repositioning) upon differentiation into epimastigotes (Sharma *et al.,* 2008). As CAP5.5 expression throughout the developmental stages in the tsetse fly is not known, one cannot speculate whether the epimastigote-like organelle configurations seen upon CAP5.5::YFP::Ty over-expression is indicative of CAP5.5's role in this differentiation process.

Organelle positioning is hypothesized to be largely mediated by microtubules, with some of the major cellular events such as kinetoplast segregation, basal body duplication, flagellar axoneme growth, mitosis and cytokinesis being shown to be dependent on microtubule polymerization/depolymerization (Robinson & Gull, 1991; Woodward & Gull, 1990; Robinson *et al.,* 1995). With regards to the kinetoplast and the nucleus, a growing body of work has established the structural, cytoskeletal linkages involved in the fidelity of kinetoplast positioning and segregation while similar information for the nucleus is lacking.

The kinetoplast is linked to the basal body complexes of the flagellum cytoskeleton by a network of filaments extending across the mitochondrial membrane called the tripartite attachment complex or TAC (Ogbadoyi *et al.,* 2003; Robinson & Gull, 1991). This structure provides the basis for kinetoplast positioning and replication that defines the various lifecycles and cell cycle stages of trypanosomes (reviewed in Gull, 2003).

On the other hand, while the nucleus has a very defined position, little is known about how such positioning, which involves nuclear migration and anchorage, is established and maintained. Known contributors to nuclear positioning in other eukaryotes have included microtubules, actin, intermediate filaments and cytoplasmic motor proteins such as dyneins and kinesins (reviewed in Dupin *et al.,* 2011; Morris *et al.,* 2003). However, in *T. brucei*, there is currently limited evidence for the presence of intermediate filament proteins, and the nucleus in both bloodstream and procyclic cells did not appear to be affected by the loss of actin (Hammarton *et al.,* 2007; Garcia-Salcedo *et al.,* 2004). Depletion of kinesins KIN13-1, KIN13-2, KIF9A or KIF9B had no noticeable effect on nuclear positioning whereas functional studies on cytoplasmic dyneins remain to be investigated (Demonchy *et al.,* 2009; Chan & Ersfeld, 2010; Chan *et al.,* 2010; Wickstead *et al.,* 2010; May *et al.,* 2012).

Ultrastructural analyses have not revealed a physical linkage between the cortical array of microtubules and the nucleus (Robinson *et al.,* 1995). My data combined with the nuclear mis-positioning observed upon depletion of CAP5.5 and *Tb*AIR9 suggest the possibility of microtubule-associated proteins playing a direct or indirect role in nuclear positioning and nuclear anchorage.

5.8 Conclusion

Given the number of kinetoplastid calpains, outnumbering those in many sequenced organisms, and the potential for kinetoplastid-specific functionality, this calpain family has only received its deserved attention in the past few years. CAP5.5 was the first kinetoplastid calpain to be characterized and in this thesis, I further characterized

this prototypical calpain *in vitro* as well as *in vivo*. I have demonstrated a cell cycle dependency on incorporation of CAP5.5 into the cytoskeleton. Given the absence of detectable proteolytic activity for CAP5.5V domain II and the non-canonical catalytic triad, a proteolytic function for CAP5.5 and other kinetoplastid calpains, all of which share similar mutations in their catalytic sites, seems unlikely. Instead, based on observations about other atypical calpains and enzymes with major substitutions in their catalytic sites, these calpains may be involved in regulatory processes (Pils & Schultz, 2004; Todd *et al.,* 2002). While this issue requires further functional studies, CAP5.5 and its paralogue are known to be essential and may be involved in various cellular processes including nuclear positioning.

6. References

Abrahamsen, M. S., Templeton, T. J., Enomoto, S., Abrahante, J. E., Zhu, G., Lancto, C. A., Deng, M., Liu, C., Widmer, G., Tzipori, S., Buck, G. A., Xu, P., Bankier, A. T., Dear, P. H., Konfortov, B. A., Spriggs, H. F., Iyer, L., Anantharaman, V., Aravind, L., and Kapur, V. (2004) Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science 304*, 441-445.

Abramoff, M. D., Magalhaes, P.J., Ram, S.J. (2004) Image Processing with ImageJ. *Biophotonics International 11*, 36-42.

Aksoy, S., Gibson, W. C., and Lehane, M. J. (2003) Interactions between tsetse and trypanosomes with implications for the control of trypanosomiasis. *Adv Parasitol 53*, 1-83.

Alsford, S., duBois, K., Horn, D., and Field, M. C. (2012) Epigenetic mechanisms, nuclear architecture and the control of gene expression in trypanosomes. *Expert Rev Mol Med 14*, e13.

Anderson, I. J., and Loftus, B. J. (2005) *Entamoeba histolytica*: observations on metabolism based on the genome sequence. *Exp Parasitol 110*, 173-177.

Andrade, H. M., Murta, S. M., Chapeaurouge, A., Perales, J., Nirde, P., and Romanha, A. J. (2008) Proteomic analysis of *Trypanosoma cruzi* resistance to Benznidazole. *J Proteome Res 7*, 2357-2367.

Arakawa, T., Philo, J. S., and Kita, Y. (2001) Kinetic and thermodynamic analysis of thermal unfolding of recombinant erythropoietin. *Biosci Biotechnol Biochem 65*, 1321-1327.

Archer, S. K., Inchaustegui, D., Queiroz, R., and Clayton, C. (2011) The cell cycle regulated transcriptome of *Trypanosoma brucei*. *PLoS One 6*, e18425.

Bacchi, C. J. (2009) Chemotherapy of human african trypanosomiasis. *Interdiscip Perspect Infect Dis 2009*, 195040.

Baines, A., and Gull, K. (2008) WCB is a C2 domain protein defining the plasma membrane - sub-pellicular microtubule corset of kinetoplastid parasites. *Protist 159*, 115-125.

Barrett, M. P., Boykin, D. W., Brun, R., and Tidwell, R. R. (2007) Human African trypanosomiasis: pharmacological re-engagement with a neglected disease. *Br J Pharmacol 152*, 1155-1171.

Barth, A. I., Nathke, I. S., and Nelson, W. J. (1997) Cadherins, catenins and APC protein: interplay between cytoskeletal complexes and signalling pathways. *Curr Opin Cell Biol 9*, 683-690.

Bastin, P., and Gull, K. (1999) Assembly and function of complex flagellar structures illustrated by the paraflagellar rod of trypanosomes. *Protist 150*, 113-123.

Bennett, V., and Stenbuck, P. J. (1979) The membrane attachment protein for spectrin is associated with band 3 in human erythrocyte membranes. *Nature 280*, 468-473.

Benz, C., Clucas, C., Mottram, J. C., and Hammarton, T. C. (2012) Cytokinesis in bloodstream stage *Trypanosoma brucei* requires a family of katanins and spastin. *PLoS One 7*, e30367.

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

Bhatt, A., Kaverina, I., Otey, C., and Huttenlocher, A. (2002) Regulation of focal complex composition and disassembly by the calcium-dependent protease calpain. *J Cell Sci 115*, 3415-3425.

Birkett, C. R., Parma, A. E., Gerke-Bonet, R., Woodward, R., and Gull, K. (1992) Isolation of cDNA clones encoding proteins of complex structures: analysis of the *Trypanosoma brucei* cytoskeleton. *Gene 110*, 65-70.

Boehmerle, W., Zhang, K., Sivula, M., Heidrich, F. M., Lee, Y., Jordt, S. E., and Ehrlich, B. E. (2007) Chronic exposure to paclitaxel diminishes phosphoinositide signalling by calpain-mediated neuronal calcium sensor-1 degradation. *Proc Natl Acad Sci U S A 104*, 11103-11108.

Borst, P. (2002) Antigenic variation and allelic exclusion. *Cell 109*, 5-8.

Bringaud, F., Riviere, L., and Coustou, V. (2006) Energy metabolism of trypanosomatids: adaptation to available carbon sources. *Mol Biochem Parasitol 149*, 1-9.

Bruce, D. (1895) *Preliminary report on the tsetse fly disease or nagana in Zululand.* Bennett and Davis, Durban.

Brun, R., and Jenni, L. (1977) A new semi-defined medium for *Trypanosoma brucei* sspp. *Acta Trop 34*, 21-33.

Buday, L. (1999) Membrane-targeting of signalling molecules by SH2/SH3 domain-containing adaptor proteins. *Biochim Biophys Acta 1422*, 187-204.

Burn, P., and Burger, M. M. (1987) The cytoskeletal protein vinculin contains transformation-sensitive, covalently bound lipid. *Science 235*, 476-479.

Campbell, R. L., and Davies, P. L. (2012) Structure-function relationships in calpains. *Biochem J 447*, 335-351.

Cazzulo, J. J., Franke de Cazzulo, B. M., Engel, J. C., and Cannata, J. J. (1985) End products and enzyme levels of aerobic glucose fermentation in trypanosomatids. *Mol Biochem Parasitol 16*, 329-343.

Chailley, B., and Boisvieux-Ulrich, E. (1985) Detection of plasma membrane cholesterol by filipin during microvillogenesis and ciliogenesis in quail oviduct. *J Histochem Cytochem 33*, 1-10.

Chan, K. T., Bennin, D. A., and Huttenlocher, A. (2010) Regulation of adhesion dynamics by calpain-mediated proteolysis of focal adhesion kinase (FAK). *J Biol Chem 285*, 11418-11426.

Chan, K. Y., and Ersfeld, K. (2010) The role of the Kinesin-13 family protein TbKif13-2 in flagellar length control of *Trypanosoma brucei. Mol Biochem Parasitol 174*, 137-140.

Chan, K. Y., Matthews, K. R., and Ersfeld, K. (2010) Functional characterisation and drug target validation of a mitotic kinesin-13 in *Trypanosoma brucei*. *PLoS Pathog 6*, e1001050.

Chen, X., Zaro, J., Shen, W.C. (2013) Fusion Protein Linkers: Effects on Production, Bioactivity, and Pharmacokinetics, In *Fusion Protein Technologies for Biopharmaceuticals: Applications and Challenges* (Schmidt, S. R., Ed.), John Wiley & Sons, Inc., Hoboken, New Jersey.

Clayton, C. E. (2002) Life without transcriptional control? From fly to man and back again. *EMBO J 21*, 1881-1888.

Cole, D. G., Diener, D. R., Himelblau, A. L., Beech, P. L., Fuster, J. C., and Rosenbaum, J. L. (1998) Chlamydomonas kinesin-II-dependent intraflagellar transport (IFT): IFT particles contain proteins required for ciliary assembly in *Caenorhabditis elegans* sensory neurons. *J Cell Biol 141*, 993-1008.

Corbalan-Garcia, S., and Gomez-Fernandez, J. C. (2010) The C2 domains of classical and novel PKCs as versatile decoders of membrane signals. *Biofactors 36*, 1-7.

Croall, D. E., and Ersfeld, K. (2007) The calpains: modular designs and functional diversity. *Genome Biol 8*, 218.

D'Arcy, A. (1994) Crystallizing proteins - a rational approach? *Acta Crystallogr D Biol Crystallogr 50*, 469-471.

de Paulo Martins, V., Okura, M., Maric, D., Engman, D. M., Vieira, M., Docampo, R., and Moreno, S. N. (2010) Acylation-dependent export of *Trypanosoma cruzi* phosphoinositide-specific phospholipase C to the outer surface of amastigotes. *J Biol Chem 285*, 30906-30917.

Demonchy, R., Blisnick, T., Deprez, C., Toutirais, G., Loussert, C., Marande, W., Grellier, P., Bastin, P., and Kohl, L. (2009) Kinesin 9 family members perform separate functions in the trypanosome flagellum. *J Cell Biol 187*, 615-622.

Denison, S. H., Orejas, M., and Arst, H. N., Jr. (1995) Signalling of ambient pH in Aspergillus involves a cysteine protease. *J Biol Chem 270*, 28519-28522.

Dennison, S. R., Dante, S., Hauss, T., Brandenburg, K., Harris, F., and Phoenix, D. A. (2005) Investigations into the membrane interactions of m-calpain domain V. *Biophys J 88*, 3008-3017.

Denny, P. W., Gokool, S., Russell, D. G., Field, M. C., and Smith, D. F. (2000) Acylation-dependent protein export in Leishmania, *J Biol Chem 275*, 11017-11025.

Docampo, R., and Moreno, S. N. (2003) Current chemotherapy of human African trypanosomiasis, *Parasitol Res 90 Supp 1*, S10-13.

Dourdin, N., Bhatt, A. K., Dutt, P., Greer, P. A., Arthur, J. S., Elce, J. S., and Huttenlocher, A. (2001) Reduced cell migration and disruption of the actin cytoskeleton in calpain-deficient embryonic fibroblasts. *J Biol Chem 276*, 48382-48388.

Dupin, I., Sakamoto, Y., and Etienne-Manneville, S. (2011) Cytoplasmic intermediate filaments mediate actin-driven positioning of the nucleus. *J Cell Sci 124*, 865-872.

Edelstein, A., Amodaj, N., Hoover, K., Vale, R., and Stuurman, N. (2010) Computer control of microscopes using micromanager. *Curr Protoc Mol Biol Chapter 14*, Unit 14 20.

Elce, J. S., Hegadorn, C., and Arthur, J. S. (1997) Autolysis, Ca^{2+} requirement, and heterodimer stability in m-calpain. *J Biol Chem 272*, 11268-11275.

Elce, J. S., Hegadorn, C., Gauthier, S., Vince, J. W., and Davies, P. L. (1995) Recombinant calpain II: improved expression systems and production of a C105A active-site mutant for crystallography. *Protein Eng 8*, 843-848.

Emmer, B. T., Nakayasu, E. S., Souther, C., Choi, H., Sobreira, T. J., Epting, C. L., Nesvizhskii, A. I., Almeida, I. C., and Engman, D. M. (2011) Global analysis of protein palmitoylation in African trypanosomes. *Eukaryot Cell 10*, 455-463.

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

Ersfeld, K., Barraclough, H., and Gull, K. (2005) Evolutionary relationships and protein domain architecture in an expanded calpain superfamily in kinetoplastid parasites. *J Mol Evol 61*, 742-757.

Ersfeld, K., and Gull, K. (2001) Targeting of cytoskeletal proteins to the flagellum of *Trypanosoma brucei. J Cell Sci 114*, 141-148.

Farr, H., and Gull, K. (2012) Cytokinesis in trypanosomes. *Cytoskeleton (Hoboken) 69*, 931-941.

Fenn, K., and Matthews, K. R. (2007) The cell biology of *Trypanosoma brucei* differentiation. *Curr Opin Microbiol 10*, 539-546.

Fernandez-Montalvan, A., Assfalg-Machleidt, I., Pfeiler, D., Fritz, H., Jochum, M., and Machleidt, W. (2004) Electrostatic interactions of domain III stabilize the inactive conformation of mu-calpain. *Biochem J 382*, 607-617.

Ferrante, A., and Allison, A. C. (1983) Alternative pathway activation of complement by African trypanosomes lacking a glycoprotein coat. *Parasite Immunol 5*, 491-498.

Figueiredo, L. M., Cross, G. A., and Janzen, C. J. (2009) Epigenetic regulation in African trypanosomes: a new kid on the block, *Nat Rev Microbiol 7*, 504-513.

Forde, R. M. (1904) Some clinical notes on a European patient in whose blood a trypanosome was observed. *J Trop Med 5*, 2.

Franco, S. J., and Huttenlocher, A. (2005) Regulating cell migration: calpains make the cut. *J Cell Sci 118*, 3829-3838.

Franz, T., Vingron, M., Boehm, T., and Dear, T. N. (1999) Capn7: a highly divergent vertebrate calpain with a novel C-terminal domain. *Mamm Genome 10*, 318-321.

Frearson, J. A., Brand, S., McElroy, S. P., Cleghorn, L. A., Smid, O., Stojanovski, L., Price, H. P., Guther, M. L., Torrie, L. S., Robinson, D. A., Hallyburton, I., Mpamhanga, C. P., Brannigan, J. A., Wilkinson, A. J., Hodgkinson, M., Hui, R., Qiu, W., Raimi, O. G., van Aalten, D. M., Brenk, R., Gilbert, I. H., Read, K. D., Fairlamb, A. H., Ferguson, M. A., Smith, D. F., and Wyatt, P. G. (2010) N-myristoyltransferase inhibitors as new leads to treat sleeping sickness. *Nature 464*, 728-732.

Fridberg, A., Buchanan, K. T., and Engman, D. M. (2007) Flagellar membrane trafficking in kinetoplastids. *Parasitol Res 100*, 205-212.

Furuya, T., Kashuba, C., Docampo, R., and Moreno, S. N. (2000) A novel phosphatidylinositol-phospholipase C of *Trypanosoma cruzi* that is lipid modified and activated during trypomastigote to amastigote differentiation. *J Biol Chem 275*, 6428-6438.

Futai, E., Kubo, T., Sorimachi, H., Suzuki, K., and Maeda, T. (2001) Molecular cloning of PalBH, a mammalian homologue of the Aspergillus atypical calpain PalB. *Biochim Biophys Acta 1517*, 316-319.

Galetovic, A., Souza, R. T., Santos, M. R., Cordero, E. M., Bastos, I. M., Santana, J. M., Ruiz, J. C., Lima, F. M., Marini, M. M., Mortara, R. A., and da Silveira, J. F. (2011) The repetitive cytoskeletal protein H49 of *Trypanosoma cruzi* is a calpain-like protein located at the flagellum attachment zone. *PLoS One 6*, e27634.

Garcia-Salcedo, J. A., Perez-Morga, D., Gijon, P., Dilbeck, V., Pays, E., and Nolan, D. P. (2004) A differential role for actin during the life cycle of *Trypanosoma brucei*. *EMBO J 23*, 780-789.

Gardner, M. J., Bishop, R., Shah, T., de Villiers, E. P., Carlton, J. M., Hall, N., Ren, Q., Paulsen, I. T., Pain, A., Berriman, M., Wilson, R. J., Sato, S., Ralph, S. A., Mann, D. J., Xiong, Z., Shallom, S. J., Weidman, J., Jiang, L., Lynn, J., Weaver, B., Shoaibi, A., Domingo, A. R., Wasawo, D., Crabtree, J., Wortman, J. R., Haas, B., Angiuoli, S. V., Creasy, T. H., Lu, C., Suh, B., Silva, J. C., Utterback, T. R., Feldblyum, T. V., Pertea, M., Allen, J., Nierman, W. C., Taracha, E. L., Salzberg, S. L., White, O. R., Fitzhugh, H. A., Morzaria, S., Venter, J. C., Fraser, C. M., and Nene, V. (2005) Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science 309*, 134-137.

Gehrig, S., and Efferth, T. (2008) Development of drug resistance in Trypanosoma brucei rhodesiense and Trypanosoma brucei gambiense. Treatment of human African trypanosomiasis with natural products (Review). *Int J Mol Med 22*, 411-419.

Ghazi-Tabatabai, S., Obita, T., Pobbati, A. V., Perisic, O., Samson, R. Y., Bell, S. D., and Williams, R. L. (2009) Evolution and assembly of ESCRTs. *Biochem Soc Trans 37*, 151-155.

Giese, V., Dallagiovanna, B., Marchini, F. K., Pavoni, D. P., Krieger, M. A., and Goldenberg, S. (2008) *Trypanosoma cruzi*: a stage-specific calpain-like protein is induced after various kinds of stress. *Mem Inst Oswaldo Cruz 103*, 598-601.

Gluenz, E., Povelones, M. L., Englund, P. T., and Gull, K. (2011) The kinetoplast duplication cycle in *Trypanosoma brucei* is orchestrated by cytoskeleton-mediated cell morphogenesis. *Mol Cell Biol 31*, 1012-1021.

Godsel, L. M., and Engman, D. M. (1999) Flagellar protein localization mediated by a calcium-myristoyl/palmitoyl switch mechanism. *EMBO J 18*, 2057-2065.

Goll, D. E., Thompson, V. F., Li, H., Wei, W., and Cong, J. (2003) The calpain system. *Physiol Rev 83*, 731-801.

Griffith, F. L. (1898) *The Petri Papyri: Hieratic Papyri from Kahun and Gurob (Principally of the Middle Kingdom).* Bernard Quaritch, London.

Gull, K. (1999) The cytoskeleton of trypanosomatid parasites. *Annu Rev Microbiol 53*, 629-655.

Gull, K. (2003) Host-parasite interactions and trypanosome morphogenesis: a flagellar pocketful of goodies. *Curr Opin Microbiol 6*, 365-370.

Gunzl, A. (2010) The pre-mRNA splicing machinery of trypanosomes: complex or simplified? *Eukaryot Cell 9*, 1159-1170.

Hammarton, T.C., Wickstead, B., McKean, P.G . (2006) Cell structure, cell division and cell cycle. In *Trypanosomes—after the genome* (Barry J.D., M. J. C., McCulloch R., Acosta-Serrano A., Ed.). Horizon Scientific Press, Norwich.

Hammarton, T. C. (2007) Cell cycle regulation in *Trypanosoma brucei*. *Mol Biochem Parasitol 153*, 1-8.

Hammarton, T. C., Clark, J., Douglas, F., Boshart, M., and Mottram, J. C. (2003) Stage-specific differences in cell cycle control in *Trypanosoma brucei* revealed by RNA interference of a mitotic cyclin. *J Biol Chem 278*, 22877-22886.

Hanna, R. A., Campbell, R. L., and Davies, P. L. (2008) Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin. *Nature 456*, 409-412.

Hata, S., Sorimachi, H., Nakagawa, K., Maeda, T., Abe, K., and Suzuki, K. (2001) Domain II of m-calpain is a Ca^{2+}-dependent cysteine protease. *FEBS Lett 501*, 111-114.

Hemphill, A., Lawson, D., and Seebeck, T. (1991) The cytoskeletal architecture of *Trypanosoma brucei. J Parasitol 77*, 603-612.

Hertz-Fowler, C., Ersfeld, K., and Gull, K. (2001) CAP5.5, a life-cycle-regulated, cytoskeleton-associated protein is a member of a novel family of calpain-related proteins in *Trypanosoma brucei. Mol Biochem Parasitol 116*, 25-34.

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

Horan, T., Wen, J., Arakawa, T., Liu, N., Brankow, D., Hu, S., Ratzkin, B., and Philo, J. S. (1995) Binding of Neu differentiation factor with the extracellular domain of Her2 and Her3. *J Biol Chem 270*, 24604-24608.

Horn, D., and Cross, G. A. (1997) Position-dependent and promoter-specific regulation of gene expression in *Trypanosoma brucei. EMBO J 16*, 7422-7431.

Hosfield, C. M., Elce, J. S., Davies, P. L., and Jia, Z. (1999) Crystal structure of calpain reveals the structural basis for Ca^{2+}-dependent protease activity and a novel mode of enzyme activation. *EMBO J 18*, 6880-6889.

Hurley, J. H., and Misra, S. (2000) Signalling and subcellular targeting by membrane-binding domains. *Annu Rev Biophys Biomol Struct 29*, 49-79.

Ito, M., Tanaka, T., Nunoki, K., Hidaka, H., and Suzuki, K. (1987) The Ca2+ -activated protease (calpain) modulates Ca^{2+}/calmodulin dependent activity of smooth muscle myosin light chain kinase. *Biochem Biophys Res Commun 145*, 1321-1328.

Jekely, G., and Friedrich, P. (1999) The evolution of the calpain family as reflected in paralogous chromosome regions. *J Mol Evol 49*, 272-281.

Joyce, P. I., Satija, R., Chen, M., and Kuwabara, P. E. (2012) The atypical calpains: evolutionary analyses and roles in *Caenorhabditis elegans* cellular degeneration. *PLoS Genet 8*, e1002602.

Kaneshiro, E. S., Matesic, D. F., and Jayasimhulu, K. (1984) Characterizations of six ethanolamine sphingophospholipids from Paramecium cells and cilia. *J Lipid Res 25*, 369-377.

Kavanagh, K. L., Jornvall, H., Persson, B., and Oppermann, U. (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families: the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell Mol Life Sci 65*, 3895-3906.

Kelley, L. A., and Sternberg, M. J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc 4*, 363-371.

Kelly, S., Reed, J., Kramer, S., Ellis, L., Webb, H., Sunter, J., Salje, J., Marinsek, N., Gull, K., Wickstead, B., and Carrington, M. (2007) Functional genomics in *Trypanosoma brucei*: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. *Mol Biochem Parasitol 154*, 103-109.

Kennedy, P. G. (2004) Human African trypanosomiasis of the CNS: current issues and challenges. *J Clin Invest 113*, 496-504.

Kennedy, P. G. (2006) Diagnostic and neuropathogenesis issues in human African trypanosomiasis. *Int J Parasitol 36*, 505-512.

Kennedy, P. G. (2008) The continuing problem of human African trypanosomiasis (sleeping sickness). *Ann Neurol 64*, 116-126.

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

Kilmartin, J. V., Wright, B., and Milstein, C. (1982) Rat monoclonal antitubulin antibodies derived by using a new nonsecreting rat cell line. *J Cell Biol 93*, 576-582.

Kirsch, J., Langosch, D., Prior, P., Littauer, U. Z., Schmitt, B., and Betz, H. (1991) The 93-kDa glycine receptor-associated protein binds to tubulin. *J Biol Chem 266*, 22242-22245.

Kohl, L., and Bastin, P. (2005) The flagellum of trypanosomes. *Int Rev Cytol 244*, 227-285.

Kohl, L., and Gull, K. (1998) Molecular architecture of the trypanosome cytoskeleton. *Mol Biochem Parasitol 93*, 1-9.

Kozminski, K. G., Johnson, K. A., Forscher, P., and Rosenbaum, J. L. (1993) A motility in the eukaryotic flagellum unrelated to flagellar beating. *Proc Natl Acad Sci U S A 90*, 5519-5523.

Kretsinger, R. H. (1997) EF-hands embrace. *Nat Struct Biol 4*, 514-516.

Lafaille, J. J., Linss, J., Krieger, M. A., Souto-Padron, T., de Souza, W., and Goldenberg, S. (1989) Structure and expression of two *Trypanosoma cruzi* genes encoding antigenic proteins bearing repetitive epitopes. *Mol Biochem Parasitol 35*, 127-136.

Lamb, A. L., Torres, A. S., O'Halloran, T. V., and Rosenzweig, A. C. (2000) Heterodimer formation between superoxide dismutase and its copper chaperone. *Biochemistry 39*, 14720-14727.

Lamour, N., Riviere, L., Coustou, V., Coombs, G. H., Barrett, M. P., and Bringaud, F. (2005) Proline metabolism in procyclic *Trypanosoma brucei* is down-regulated in the presence of glucose. *J Biol Chem 280*, 11902-11910.

Larsen, A. K., De Veyra, T., Jia, Z., Wells, A., Dutt, P., and Elce, J. S. (2004) Expression of human, mouse, and rat m-calpains in *Escherichia coli* and in murine fibroblasts. *Protein Expr Purif 33*, 246-255.

Lebart, M. C., and Benyamin, Y. (2006) Calpain involvement in the remodelling of cytoskeletal anchorage complexes. *FEBS J 273*, 3415-3426.

Lewis, E. A., and Langridge, W. P. (1947) Developmental forms of *Trypanosoma brucei* in the saliva of *Glossina pallidipes* and *Glossina austeni*. *Ann Trop Med Parasitol 41*, 6-13.

Li, Z. (2012) Regulation of the cell division cycle in *Trypanosoma brucei. Eukaryot Cell 11*, 1180-1190.

Liu, W., Apagyi, K., McLeavy, L., and Ersfeld, K. (2010) Expression and cellular localisation of calpain-like proteins in *Trypanosoma brucei. Mol Biochem Parasitol 169*, 20-26.

Lopez-Mirabal, H. R., and Winther, J. R. (2008) Redox characteristics of the eukaryotic cytosol. *Biochim Biophys Acta 1783*, 629-640.

Maretzki, D., Mariani, M., and Lutz, H. U. (1990) Fatty acid acylation of membrane skeletal proteins in human erythrocytes. *FEBS Lett 259*, 305-310.

Mariani, M., Maretzki, D., and Lutz, H. U. (1993) A tightly membrane-associated subpopulation of spectrin is 3H-palmitoylated. *J Biol Chem 268*, 12996-13001.

Maric, D., McGwire, B. S., Buchanan, K. T., Olson, C. L., Emmer, B. T., Epting, C. L., and Engman, D. M. (2011) Molecular determinants of ciliary membrane localization of *Trypanosoma cruzi* flagellar calcium-binding protein. *J Biol Chem 286*, 33109-33117.

Marr, J. J., and Berens, R. L. (1977) Regulation of aerobic fermentation in protozoans. VI Comparative biochemistry of pathogenic and nonpathogenic protozoans. *Acta Trop 34*, 143-155.

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

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

Matthews, K. R., and Gull, K. (1994a) Evidence for an interplay between cell cycle progression and the initiation of differentiation between life cycle forms of African trypanosomes. *J Cell Biol 125*, 1147-1156.

Matthews, K. R., and Gull, K. (1994b) Cycles within cycles: the interplay between differentiation and cell division in *Trypanosoma brucei. Parasitol Today 10*, 473-476.

Matthews, K. R., Sherwin, T., and Gull, K. (1995) Mitochondrial genome repositioning during the differentiation of the African trypanosome between life cycle forms is microtubule mediated. *J Cell Sci 108 (Pt 6)*, 2231-2239.

May, S. F., Peacock, L., Almeida Costa, C. I., Gibson, W. C., Tetley, L., Robinson, D. R., and Hammarton, T. C. (2012) The *Trypanosoma brucei* AIR9-like protein is cytoskeleton-associated and is required for nucleus positioning and accurate cleavage furrow placement. *Mol Microbiol 84*, 77-92.

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

McKean, P. G. (2003) Coordination of cell cycle and cytokinesis in *Trypanosoma brucei*. *Curr Opin Microbiol 6*, 600-607.

Mills, E., Price, H. P., Johner, A., Emerson, J. E., and Smith, D. F. (2007) Kinetoplastid PPEF phosphatases: dual acylated proteins expressed in the endomembrane system of Leishmania. *Mol Biochem Parasitol 152*, 22-34.

Moldoveanu, T., Campbell, R. L., Cuerrier, D., and Davies, P. L. (2004a) Crystal structures of calpain-E64 and -leupeptin inhibitor complexes reveal mobile loops gating the active site. *J Mol Biol 343*, 1313-1326.

Moldoveanu, T., Hosfield, C. M., Lim, D., Elce, J. S., Jia, Z., and Davies, P. L. (2002) A Ca^{2+} switch aligns the active site of calpain, *Cell 108*, 649-660.

Moldoveanu, T., Jia, Z., and Davies, P. L. (2004b) Calpain activation by cooperative Ca^{2+} binding at two non-EF-hand sites. *J Biol Chem 279*, 6106-6114.

Morris, N. R. (2003) Nuclear positioning: the means is at the ends. *Curr Opin Cell Biol 15*, 54-59.

Morrow, J. S., Cianci, C. D., Ardito, T., Mann, A. S., and Kashgarian, M. (1989) Ankyrin links fodrin to the alpha subunit of Na,K-ATPase in Madin-Darby canine kidney cells and in intact renal tubule cells. *J Cell Biol 108*, 455-465.

Muller, N., Hemphill, A., Imboden, M., Duvallet, G., Dwinger, R. H., and Seebeck, T. (1992) Identification and characterization of two repetitive non-variable antigens from African trypanosomes which are recognized early during infection. *Parasitology 104 Pt 1*, 111-120.

Mumby, S. M., Kleuss, C., and Gilman, A. G. (1994) Receptor regulation of G-protein palmitoylation. *Proc Natl Acad Sci U S A 91*, 2800-2804.

Nasser, M. I., and Landfear, S. M. (2004) Sequences required for the flagellar targeting of an integral membrane protein. *Mol Biochem Parasitol 135*, 89-100.

Neugebauer, J. M. (1990) Detergents: an overview. *Methods Enzymol 182*, 239-253.

Niebuhr, K., Ebel, F., Frank, R., Reinhard, M., Domann, E., Carl, U. D., Walter, U., Gertler, F. B., Wehland, J., and Chakraborty, T. (1997) A novel proline-rich motif present in ActA of *Listeria monocytogenes* and cytoskeletal proteins is the ligand for the EVH1 domain, a protein module present in the Ena/VASP family. *EMBO J 16*, 5433-5444.

Ogbadoyi, E. O., Robinson, D. R., and Gull, K. (2003) A high-order trans-membrane structural linkage is responsible for mitochondrial genome positioning and segregation by flagellar basal bodies in trypanosomes. *Mol Biol Cell 14*, 1769-1779.

Ohno, S. (1970) *Evolution by gene duplication*. Spronger-Verlag, New York.

Ohno, S., Emori, Y., Imajoh, S., Kawasaki, H., Kisaragi, M., and Suzuki, K. (1984) Evolutionary origin of a calcium-dependent protease by fusion of genes for a thiol protease and a calcium-binding protein? *Nature 312*, 566-570.

Okura, M., Fang, J., Salto, M. L., Singer, R. S., Docampo, R., and Moreno, S. N. (2005) A lipid-modified phosphoinositide-specific phospholipase C (TcPI-PLC) is involved in differentiation of trypomastigotes to amastigotes of *Trypanosoma cruzi*. *J Biol Chem 280*, 16235-16243.

Olego-Fernandez, S. (2010) An Expanded Calpain-like Family of Proteins in Kinetoplastid Parasites, In *Sir William Dunn School of Pathology*. University of Oxford, Oxford.

Olego-Fernandez, S., Vaughan, S., Shaw, M. K., Gull, K., and Ginger, M. L. (2009) Cell morphogenesis of *Trypanosoma brucei* requires the paralogous, differentially expressed calpain-related proteins CAP5.5 and CAP5.5V. *Protist 160*, 576-590.

Ono, Y., and Sorimachi, H. (2012) Calpains: an elaborate proteolytic system. *Biochim Biophys Acta 1824*, 224-236.

Osako, Y., Maemoto, Y., Tanaka, R., Suzuki, H., Shibata, H., and Maki, M. (2010) Autolytic activity of human calpain 7 is enhanced by ESCRT-III-related protein IST1 through MIT-MIM interaction. *FEBS J 277*, 4412-4426.

Ouellette, M., and Papadopoulou, B. (2009) Coordinated gene expression by post-transcriptional regulons in African trypanosomes. *J Biol 8*, 100.

Pal, G. P., Elce, J. S., and Jia, Z. (2001) Dissociation and aggregation of calpain in the presence of calcium. *J Biol Chem 276*, 47233-47238.

Palenchar, J. B., and Bellofatto, V. (2006) Gene transcription in trypanosomes. *Mol Biochem Parasitol 146*, 135-141.

Pan, S. H., and Malcolm, B. A. (2000) Reduced background expression and improved plasmid stability with pET vectors in BL21 (DE3). *Biotechniques 29*, 1234-1238.

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

Pazour, G. J., Dickert, B. L., and Witman, G. B. (1999) The DHC1b (DHC2) isoform of cytoplasmic dynein is required for flagellar assembly. *J Cell Biol 144*, 473-481.

Pentreath, V. W. (1995) Royal Society of Tropical Medicine and Hygiene Meeting at Manson House, London, 19 May 1994. Trypanosomiasis and the nervous system. Pathology and immunology, *Trans R Soc Trop Med Hyg 89*, 9-15.

Pietsch, M., Chua, K. C., and Abell, A. D. (2010) Calpains: attractive targets for the development of synthetic inhibitors. *Curr Top Med Chem 10*, 270-293.

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

Ploubidou, A., Robinson, D. R., Docherty, R. C., Ogbadoyi, E. O., and Gull, K. (1999) Evidence for novel cell cycle checkpoints in trypanosomes: kinetoplast segregation and cytokinesis in the absence of mitosis. *J Cell Sci 112 (Pt 24)*, 4641-4650.

Poon, S. K., Peacock, L., Gibson, W., Gull, K., and Kelly, S. (2012) A modular and optimized single marker system for generating *Trypanosoma brucei* cell lines expressing T7 RNA polymerase and the tetracycline repressor. *Open Biol 2*, 110037.

Portman, N., and Gull, K. (2012) Proteomics and the *Trypanosoma brucei* cytoskeleton: advances and opportunities, *Parasitology 139*, 1168-1177.

Price, H. P., Menon, M. R., Panethymitaki, C., Goulding, D., McKean, P. G., and Smith, D. F. (2003) Myristoyl-CoA:protein N-myristoyltransferase, an essential enzyme and potential drug target in kinetoplastid parasites. *J Biol Chem 278*, 7206-7214.

Ravulapalli, R., Campbell, R. L., Gauthier, S. Y., Dhe-Paganon, S., and Davies, P. L. (2009) Distinguishing between calpain heterodimerization and homodimerization. *FEBS J 276*, 973-982.

Ravulapalli, R., Diaz, B. G., Campbell, R. L., and Davies, P. L. (2005) Homodimerization of calpain 3 penta-EF-hand domain. *Biochem J 388*, 585-591.

Ren, R., Mayer, B. J., Cicchetti, P., and Baltimore, D. (1993) Identification of a ten-amino acid proline-rich SH3 binding site. *Science 259*, 1157-1161.

Resh, M. D. (1996) Regulation of cellular signalling by fatty acid acylation and prenylation of signal transduction proteins. *Cell Signal 8*, 403-412.

Resh, M. D. (1999) Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim Biophys Acta 1451*, 1-16.

Resh, M. D. (2006) Trafficking and signalling by fatty-acylated and prenylated proteins. *Nat Chem Biol 2*, 584-590.

Reuner, B., Vassella, E., Yutzy, B., and Boshart, M. (1997) Cell density triggers slender to stumpy differentiation of *Trypanosoma brucei* bloodstream forms in culture. *Mol Biochem Parasitol 90*, 269-280.

Rizo, J., and Sudhof, T. C. (1998) C2-domains, structure and function of a universal Ca2+-binding domain. *J Biol Chem 273*, 15879-15882.

Robertson, M. (1913) Notes on the polymorphism of *Trypanosoma gambiense* in the blood and its relation to the exogenous cycle in *Glossina palpalis*. *Phil. Trans. R. Soc. Lond. (Biol.) 203*, 161-184.

Robinson, D. R., and Gull, K. (1991) Basal body movements as a mechanism for mitochondrial genome segregation in the trypanosome cell cycle. *Nature 352*, 731-733.

Robinson, D. R., Sherwin, T., Ploubidou, A., Byard, E. H., and Gull, K. (1995) Microtubule polarity and dynamics in the control of organelle positioning, segregation, and cytokinesis in the trypanosome cell cycle. *J Cell Biol 128*, 1163-1172.

Rosenbaum, J. L., Cole, D. G., and Diener, D. R. (1999) Intraflagellar transport: the eyes have it. *J Cell Biol 144*, 385-388.

Russo, I., Oksman, A., and Goldberg, D. E. (2009) Fatty acid acylation regulates trafficking of the unusual *Plasmodium falciparum* calpain to the nucleolus. *Mol Microbiol 72*, 229-245.

Sarkany, Z., and Polgar, L. (2003) The unusual catalytic triad of poliovirus protease 3C. *Biochem J 42*, 516-522.

Sbicego, S., Vassella, E., Kurath, U., Blum, B., and Roditi, I. (1999) The use of transgenic *Trypanosoma brucei* to identify compounds inducing the differentiation of bloodstream forms to procyclic forms. *Mol Biochem Parasitol 104*, 311-322.

Schneider, A., Eichenberger, W., and Seebeck, T. (1988) A microtubule-binding protein of *Trypanosoma brucei* which contains covalently bound fatty acid. *J Biol Chem 263*, 6472-6475.

Schoenwaelder, S. M., and Burridge, K. (1999) Bidirectional signalling between the cytoskeleton and integrins. *Curr Opin Cell Biol 11*, 274-286.

Sharma, R., Gluenz, E., Peacock, L., Gibson, W., Gull, K., and Carrington, M. (2009) The heart of darkness: growth and form of *Trypanosoma brucei* in the tsetse fly. *Trends Parasitol 25*, 517-524.

Sharma, R., Peacock, L., Gluenz, E., Gull, K., Gibson, W., and Carrington, M. (2008) Asymmetric cell division as a route to reduction in cell length and change in cell morphology in trypanosomes. *Protist 159*, 137-151.

Sherwin, T., and Gull, K. (1989) The cell division cycle of *Trypanosoma brucei brucei*: timing of event markers and cytoskeletal modulations. *Philos Trans R Soc Lond B Biol Sci 323*, 573-588.

Sherwin, T., and Gull, K. (1989) Visualization of detyrosination along single microtubules reveals novel mechanisms of assembly during cytoskeletal duplication in trypanosomes. *Cell 57*, 211-221.

Sherwin, T., Schneider, A., Sasse, R., Seebeck, T., and Gull, K. (1987) Distinct localization and cell cycle dependence of COOH terminally tyrosinolated alpha-tubulin in the microtubules of *Trypanosoma brucei brucei. J Cell Biol 104*, 439-446.

Smythe, C. V. (1936) The reaction of iodoacetate and of iodoacetamide with various sulfhydryl groups, with urease, and with yeast preparations. *J Biol Chem 114,* 601-612. .

Snapp, E. L., and Landfear, S. M. (1997) Cytoskeletal association is important for differential targeting of glucose transporter isoforms in Leishmania. *J Cell Biol 139*, 1775-1783.

Sorimachi, H., Hata, S., and Ono, Y. (2011a) Calpain chronicle-an enzyme family under multidisciplinary characterization. *Proc Jpn Acad Ser B Phys Biol Sci 87*, 287-327.

Sorimachi, H., Hata, S., and Ono, Y. (2011b) Impact of genetic insights into calpain biology. *Biochem J 150*, 23-37.

Staufenbiel, M., and Lazarides, E. (1986) Ankyrin is fatty acid acylated in erythrocytes. *Proc Natl Acad Sci U S A 83*, 318-322.

Stegemann, J., Ventzki, R., Schrodel, A., and de Marco, A. (2005) Comparative analysis of protein aggregates by blue native electrophoresis and subsequent sodium dodecyl sulfate-polyacrylamide gel electrophoresis in a three-dimensional geometry gel. *Proteomics 5*, 2002-2009.

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

Steverding, D. (2010) The development of drugs for treatment of sleeping sickness: a historical review, *Parasit Vectors 3*, 15.

Strobl, S., Fernandez-Catalan, C., Braun, M., Huber, R., Masumoto, H., Nakagawa, K., Irie, A., Sorimachi, H., Bourenkow, G., Bartunik, H., Suzuki, K., and Bode, W. (2000) The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. *Proc Natl Acad Sci U S A 97*, 588-592.

Subramaniam, C., Veazey, P., Redmond, S., Hayes-Sinclair, J., Chambers, E., Carrington, M., Gull, K., Matthews, K., Horn, D., and Field, M. C. (2006) Chromosome-wide analysis of gene function by RNA interference in the African trypanosome. *Eukaryot Cell 5*, 1539-1549.

Suzuki, K., Imajoh, S., Emori, Y., Kawasaki, H., Minami, Y., and Ohno, S. (1987) Calcium-activated neutral protease and its endogenous inhibitor. Activation at the cell membrane and biological function. *FEBS Lett 220*, 271-277.

Tetley, L. (1986) Freeze-fracture studies on the surface membranes of pleomorphic bloodstream and in vitro transformed procyclic *Trypanosoma brucei*. *Acta Trop 43*, 307-317.

Tetley, L., and Vickerman, K. (1985) Differentiation in *Trypanosoma brucei*: host-parasite cell junctions and their persistence during acquisition of the variable antigen coat. *J Cell Sci 74*, 1-19.

Todd, A. E., Orengo, C. A., and Thornton, J. M. (2001) Evolution of function in protein superfamilies, from a structural perspective. *J Mol Biol 307*, 1113-1143.

Todd, A. E., Orengo, C. A., and Thornton, J. M. (2002) Sequence and structural differences between enzyme and nonenzyme homologs. *Structure 10*, 1435-1451.

Tonami, K., Kurihara, Y., Aburatani, H., Uchijima, Y., Asano, T., and Kurihara, H. (2007) Calpain 6 is involved in microtubule stabilization and cytoskeletal organization. *Mol Cell Biol 27*, 2548-2561.

Towler, D. A., Eubanks, S. R., Towery, D. S., Adams, S. P., and Glaser, L. (1987) Amino-terminal processing of proteins by N-myristoylation. Substrate specificity of N-myristoyl transferase. *J Biol Chem 262*, 1030-1036.

Tull, D., Heng, J., Gooley, P. R., Naderer, T., and McConville, M. J. (2012) Acylation-dependent and-independent membrane targeting and distinct functions of small myristoylated proteins (SMPs) in *Leishmania major. Int J Parasitol 42*, 239-247.

Tull, D., Naderer, T., Spurck, T., Mertens, H. D., Heng, J., McFadden, G. I., Gooley, P. R., and McConville, M. J. (2010) Membrane protein SMP-1 is required for normal flagellum function in Leishmania. *J Cell Sci123*, 544-554.

Tull, D., Vince, J. E., Callaghan, J. M., Naderer, T., Spurck, T., McFadden, G. I., Currie, G., Ferguson, K., Bacic, A., and McConville, M. J. (2004) SMP-1, a member of a new family of small myristoylated proteins in kinetoplastid parasites, is targeted to the flagellum membrane in Leishmania. *Mol Biol Cell 15*, 4775-4786.

Turner, C. M., Barry, J. D., and Vickerman, K. (1988) Loss of variable antigen during transformation of *Trypanosoma brucei rhodesiense* from bloodstream to procyclic forms in the tsetse fly. *Parasitol Res 74*, 507-511.

Tyler, K. M., Fridberg, A., Toriello, K. M., Olson, C. L., Cieslak, J. A., Hazlett, T. L., and Engman, D. M. (2009) Flagellar membrane localization via association with lipid rafts. *J Cell Sci 122*, 859-866.

Ullu, E., Matthews, K. R., and Tschudi, C. (1993) Temporal order of RNA-processing reactions in trypanosomes: rapid trans splicing precedes polyadenylation of newly synthesized tubulin transcripts. *Mol Cell Biol 13*, 720-725.

Urwyler, S., Studer, E., Renggli, C. K., and Roditi, I. (2007) A family of stage-specific alanine-rich proteins on the surface of epimastigote forms of *Trypanosoma brucei*. *Mol Microbiol 63*, 218-228.

Van Den Abbeele, J., Claes, Y., van Bockstaele, D., Le Ray, D., and Coosemans, M. (1999) *Trypanosoma brucei* spp. development in the tsetse fly: characterization of the post-mesocyclic stages in the foregut and proboscis. *Parasitology 118 (Pt 5)*, 469-478.

Vassella, E., Straesser, K., and Boshart, M. (1997) A mitochondrion-specific dye for multicolour fluorescent imaging of *Trypanosoma brucei*. *Mol Biochem Parasitol 90*, 381-385.

Vaughan, S., and Gull, K. (2003) The trypanosome flagellum. *J Cell Sci 116*, 757-759.

Vaughan, S., and Gull, K. (2008) The structural mechanics of cell division in *Trypanosoma brucei*. *Biochem Soc Trans 36*, 421-424.

Vedrenne, C., Giroud, C., Robinson, D. R., Besteiro, S., Bosc, C., Bringaud, F., and Baltz, T. (2002) Two related subpellicular cytoskeleton-associated proteins in *Trypanosoma brucei* stabilize microtubules. *Mol Biol Cell 13*, 1058-1070.

Vickerman, K. P., T. M. (1976) Comparative cell biology of the kinetoplastid flagellates. In *Biology of the Kinetoplastida*, pp 35-130, Academic Press, New York.

Wang, C., Barry, J. K., Min, Z., Tordsen, G., Rao, A. G., and Olsen, O. A. (2003) The calpain domain of the maize DEK1 protein contains the conserved catalytic triad and functions as a cysteine proteinase. *J Biol Chem 278*, 34467-34474.

Wang, C. C. (1995) Molecular mechanisms and therapeutic approaches to the treatment of African trypanosomiasis. *Annu Rev Pharmacol Toxicol 35*, 93-127.

Wickstead, B., Carrington, J. T., Gluenz, E., and Gull, K. (2010) The expanded Kinesin-13 repertoire of trypanosomes contains only one mitotic Kinesin indicating multiple extra-nuclear roles. *PLoS One 5*, e15020.

Wilson, C. A., Kreychman, J., and Gerstein, M. (2000) Assessing annotation transfer for genomics: quantifying the relations between protein sequence, structure and function through traditional and probabilistic scores. *J Mol Biol 297*, 233-249.

Wirtz, E., and Clayton, C. (1995) Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. *Science 268*, 1179-1183.

Woodward, R., and Gull, K. (1990) Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei. J Cell Sci 95 (Pt 1)*, 49-57.

Yorikawa, C., Takaya, E., Osako, Y., Tanaka, R., Terasawa, Y., Hamakubo, T., Mochizuki, Y., Iwanari, H., Kodama, T., Maeda, T., Hitomi, K., Shibata, H., and Maki, M. (2008) Human calpain 7/PalBH associates with a subset of ESCRT-III-related proteins in its N-terminal region and partly localizes to endocytic membrane compartments. *Biochem J 143*, 731-745.

Yoshizawa, T., Sorimachi, H., Tomioka, S., Ishiura, S., and Suzuki, K. (1995a) Calpain dissociates into subunits in the presence of calcium ions. *Biochem Biophys Res Commun 208*, 376-383.

Yoshizawa, T., Sorimachi, H., Tomioka, S., Ishiura, S., and Suzuki, K. (1995b) A catalytic subunit of calpain possesses full proteolytic activity. *FEBS Lett 358*, 101-103.

Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schreiber, S. L. (1994) Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell 76*, 933-945.

Yu, J. W., Jeffrey, P. D., and Shi, Y. (2009) Mechanism of procaspase-8 activation by c-FLIPL. *Proc Natl Acad Sci U S A 106*, 8169-8174.

Zettl, M., Adrain, C., Strisovsky, K., Lastun, V., and Freeman, M. (2011) Rhomboid family pseudoproteases use the ER quality control machinery to regulate intercellular signalling. *Cell 145*, 79-91.

Zhang, Y. (2009) Protein structure prediction: when is it useful? *Curr Opin Struct Biol 19*, 145-155.

Zhao, S., Liang, Z., Demko, V., Wilson, R., Johansen, W., Olsen, O. A., and Shalchian-Tabrizi, K. (2012) Massive expansion of the calpain gene family in unicellular eukaryotes. *BMC Evol Biol 12*, 193.

Ziegelbauer, K., and Overath, P. (1993) Organization of two invariant surface glycoproteins in the surface coat of *Trypanosoma brucei*. *Infect Immun 61*, 4540-4545.

Zulauf, M., and D'Arcy, A. (1992) Light Scattering of Proteins as a Criterion for Crystallization. *Journal of Crystal Growth 122*, 4.