MUTATIONAL ANALYSIS OF *T. BRUCEI* COMPONENTS OF MOTILE FLAGELLA (TbCMF) GENES IN THE AFRICAN TRYPANOSOME

by
Julie D. Hare

A Paper Presented to the Faculty of Mount Holyoke College in Partial Fulfillment of the Requirements for the Degree of Bachelors of Arts with Honor

Department of Biological Sciences

South Hadley, MA 01075

May, 2007

This paper was prepared under the direction of
Professor Amy Springer

for eight credits.

ACKNOWLEDGEMENTS

Firstly, I would like to thank the UCLA Hill lab team: Dr. Kent Hill
for the opportunity to contribute to his project
and the remaining Hill lab members Desiree Baron,
Katherine Palston, Z. Pius Kabututu and Lena Perna
for their knowledge and support.

Much gratitude to Dr. Michele Klingbeil for her wisdom
and use of her lab.

Also thanks to my committee members, Sharon Stranford and Sheila Browne, two outstanding professors, both of whom I have had the pleasure of being taught by.

And last, but certainly not least, I would like to thank the Trypanosome lab.
To my advisor, Amy Springer, thank you. Your undying patience was never left unnoticed; you have taught me so much.
And to Noël, Lindsey and Hanna, my lab mates, lab would not have been remotely the same without your friendship and knowledge.

ABSTRACT

Flagella and cilia play significant roles in mobility and environmental sensing in microbes as well as organ and gamete function in more complex organisms. The African trypanosome, *Trypanosoma brucei,* is a model organism for laboratory study of the components of the flagellum. Its single flagellum bears the conserved flagellar structure seen in many organisms. Little is known about the identities and interactions of the protein groups comprising the flagella despite familiarity of its shape. Understanding how molecular activity gives rise to coordinated movement (wave propagation) and environmental sensing is crucial to understanding and treating the diseases caused by African trypanosomes as well as diseases that stem from ciliary malfunctions in humans.

In collaboration with the Kent Hill research team at UCLA I am seeking to characterize a protein in a family called the *Trypanosoma brucei* Components of Motile Flagella (TbCMF). This family has been defined by the Hill research team through bioinformatics comparison of homologues present in species with motile flagella (*H. sapiens*, *M. musculus*, *D. melanogaster*, *C. reinhardtii*, and *C. elegans*).

RNAi will be used to characterize TbCMF 63. The predicted sequence of this gene contains a calcium binding domain. I am performing site specific mutagenesis on TbCMF 63 to perturb the function of the calcium binding protein in order to characterize its function in motility.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

Trypanosomiasis

African sleeping sickness, or African trypanosomiasis, is endemic to sub-Saharan Africa and the Americas affecting large human populations as well as other mammals. There are two variations of the human disease caused by two different sub species of the infecting organism, *Trypanosoma brucei*, the African trypanosome. The trypanosome is a parasitic protozoan (García et al., 2006) and has several genera, including *Trypanosoma* and *Leishmania* (Croft et al., 2005). These organisms are common to Central and South America as well as Africa. In these areas trypanosomes cause a chronic illness called Chagas. In both Africa and the Americas trypanosomiasis continues to be a contemporary problem (Croft et al., 2005). This paper will focus on human African trypanosomiasis (HAT) and the African species of trypanosomes.

HAT is considered a re-emerging disease as incidences of infection have continued to increase over the last several decades (Kennedy, 2004). Collectively, HAT infects an estimated 50,000 to 70,000 new people each year in the sub-Saharan region of Africa (Deborggraeve et al., 2006). The Democratic Republic of Congo, the Republic of Congo, Angola, Central African Republic and Southern Sudan have the most seriously infected populations (Chappuis et al., 2005). In the Democratic Republic of Congo (DRC) HAT incidences rose tremendously during the 1990's; infections began occurring in urban areas as well (Lutumba et al., 2005). Due to the combined efforts of the WHO, national control programs,

nongovernmental organizations and the private sector reported cases of trypanosomiasis have leveled off and begun to decline (Pepin 2007).

Below, Figure 1 shows the African distribution and level of severity of HAT infection.

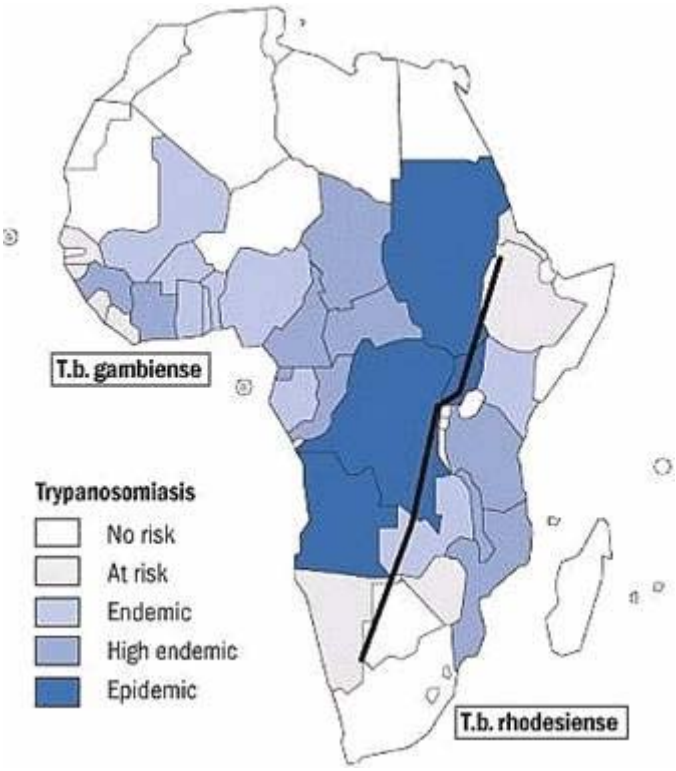


Figure 1 Map of Africa showing East and West HAT affected countries and their levels of endemic severity in 1999. Map courtesy of the World Health Organization, taken from *WHO Reports on Global Surveillance of Epidemic-Prone Infectious Disease- African trypanosomiasis*: http://www.who.int/csr/resources/publications/CSR_ISR_2000_1tryps/en/

Socio-economic perturbations, particularly war, have contributed to the disruption of disease treatment and control (Kennedy, 2004), increasing infection rates over the last several decades. Also, other mammals contribute to the continuation of infection of human populations (Deborggraeve et al., 2006).

Mammals, wild and domestic, can become infected with trypanosomes as well as humans, serving as reservoirs for human infecting trypanosomes that are picked up by the tsetse (Kennedy, 2004).

People are also adversely affected on an economic level by trypanosome-infected animals. The blue regions shown in the map in Figure 1 also correspond to more the heavily forested region of Africa. These areas breed the insect vector for trypanosomes while simultaneously provide shelter and work for people. Sub-Saharan Africa is mostly rural communities, many of which rely on farming and cattle raising for their livelihoods (Chappuis et al., 2005). Here livestock populations are often decimated by Animal African Trypanosomiasis (AAT), (Stich et al., 2006) called Nagana, the local name for AAT in cattle.

The Trypanosome

As mentioned above, trypanosomes fall under more than one genus. Each of these genera contain several species of trypanosome, each species giving rise to a different form of the disease. Discussion will be limited to species in the *Trypanosoma* genus. East African trypanosomiasis is caused by *T. brucei rhodesiense* and the West African form is due to *T. brucei gambiense* infection (García et al., 2006). It is these disease variants that are collectively referred to as African sleeping sickness. There is a difference in virulence between these two subsets of HAT, *T.b. rhodesiense* producing the faster progressing form of the disease while *T.b. gambiense* is a slower, more chronic form (García et al., 2006). Like many vector-born diseases, trypanosomes are carried and transmitted by

arthropods. An SEM of a single trypanosome is pictured below in Figure 2. It is shown among some red blood cells. In Africa it is the tsetse fly, *Glossina* spp. that spreads infection of the various *Trypanosoma* spp. (Matthews 2005).

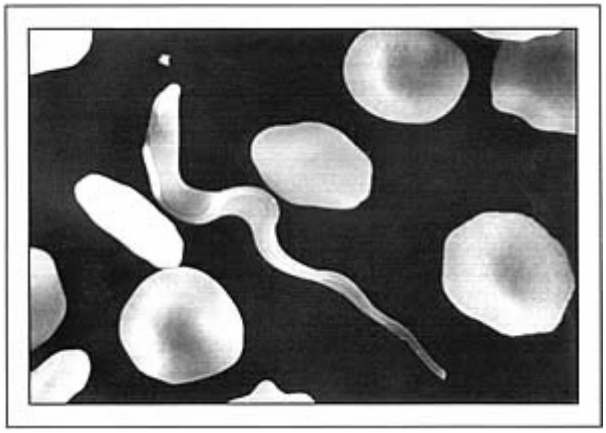


Figure 2 A trypanosome among red blood cells.
Photo courtesy of: http://sdrc.lib.uiowa.edu/preslectures/donelson90/index.html

The precise stage of HAT infection in a given patient can be difficult to diagnose. Both varieties of HAT progress in two stages and are always fatal if left untreated (Courtioux et al., 2006). Second stage HAT symptoms can appear over a period of months to years depending on the infecting trypanosome species (Vincendeau & Bouteille, 2006). The disease itself presents early on with common flu-like symptoms and is often misdiagnosed as other diseases in some patients, such as malaria, enteric fever, tuberculous meningitis and HIV infection. These diseases present similar symptoms as trypanosome infection or can occur in conjunction with it. Hence, many HAT cases have been initially overlooked (Chappuis et al., 2005).

Stage I, the hemolymphatic or early stage, begins a few days after a tsetse fly inoculates a host with trypanosomes (Vincendeau & Bouteille, 2006). The first symptom is local inflammation that occurs within 5 days at the site of the fly bite, called a trypanosomal chancre (Stich et al., 2006). The inflammation arises from the multiplying population of trypanosomes. Eventually, the infection spreads to the lymphatic system and other organs via the blood stream (Chappuis et al., 2005). The host can experience fever that coincides with fluctuating parasitemia. Also, skin lesions or rash, edema and anemia are common as the trypanosomes multiply in the blood and lymph (Courtioux et al., 2006). The progression of stage I to stage II occurs when the infecting trypanosomes invade the central nervous system (Courtioux et al., 2006). It is in this stage that various neurological symptoms begin. For this reason stage II infection is also known as the meningoencephalitic stage (Vincendeau & Bouteille, 2006). The transition of stage I to stage II happens gradually over time, as long as months or years with *T.b. gambiense*. Some of the clinical symptoms marking the second stage are altered mental status, sleep pattern disruptions, tremors, impotence, infertility and rash. These are accompanied by stage I symptoms such as fevers and skin lesions that often persist (Chappuis et al., 2005). Left untreated the patient eventually dies in a cachexiatic (immobilized by atrophy and weakness), demented state as the central nervous system demylenates and atrophies. Often the patient dies in a coma (Vincendeau, 2006).

Drug Therapies for HAT Patients

Presently there are only a few options for therapeutic drugs available to treat the two stages of trypanosomiasis (Stich et al., 2006). The newest chemotherapy used to treat late stage *gambiense* disease is eflornithine (Burri & Brun 2003). The second newest drug is melarsoprol. It was developed over 50 years ago and is still used today as the main treatment for second stage HAT (Stich et al., 2006). Eflornithine is now used primarily as a back-up in case of failure with the primary treatment for stage II, an arsenical called melarsoprol (Burri & Brun 2003). Other drugs in use include nifurtimox, a drug initially used for Chagas disease (Bisser et al., 2007), and other arsenical compounds such as pentamide and suramin (Nok 2003). The chemical structures of melarsoprol, nifurtimox and eflornithine are shown in Figure 3.

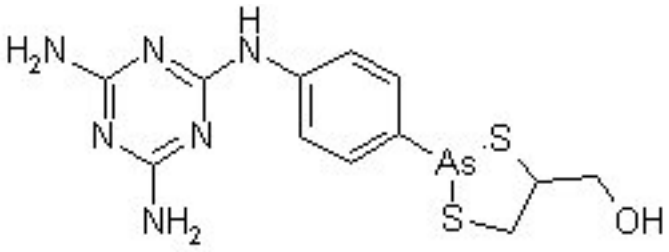
Melarsoprol has remained the top drug used for stage II HAT (Nok 2003) despite the serious toxicity associated with the treatment (Priotto et al., 2006). It is an arsenical derivative that acts as an effective trypanocidal molecule, passing through the blood-brain barrier (BBB) and causing the trypanosomes to lyse in the host's serum (Kennedy 2004). Its cytotoxic effects also work on the patient's body, inflaming and damaging tissues and organs, making this therapy intensely poisonous (Stich et al., 2006). Roughly 10% of patients receiving melarsoprol experience reactive fatal encephalopathy, an alteration of brain function and structure (Priotto, 2006).

Nifurtimox is a 5-nitrofuran molecule that is being used for *gambiense* patients that have developed resistance to melarsoprol (Pepin et al., 1992). It induces oxidative stress to the infecting trypanosomes as a result of the nitrofuran reduction (Bisser et al., 2007). Consequently, surrounding host tissue is also damaged, causing neurological and gastrointestinal disorders that worsen with increased duration of use (Priotto et al., 2006). Nifurtimox has been used with the *gambiense* form of HAT. Some patients experience relapse of trypanosomiasis. The toxicity experienced in all patients, fatal in some cases, make nifurtimox a less desirable stand alone drug (Pepin et al., 1992). A recent study showed nifurtimox combined with eflornithine (discussed below) showed evidence of being effective and less harmful than the traditional melarsoprol therapy (Priotto et al., 2006).

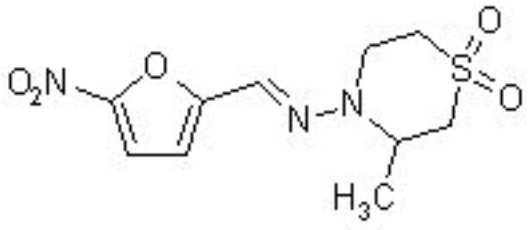
Eflornithine is also called DFMO, an abbreviation for diethylfluoromethylornithine. This drug was initially developed as a cancer treatment over 20 years ago (Kennedy 2004). It was not approved by the US FDA for treatment of West African trypanosomiasis until 1990 despite its usage since the 1980's (Burri & Brun, 2003). Eflornithine irreversibly inhibits ornithine decarboxylase (ODC), a major enzyme involved in polyamine biosynthesis (Kennedy 2004). Nucleic acid synthesis and protein synthesis depend on the production and availability of polyamines as building blocks for larger biomolecules. Blocking polyamine production effectively halts the cells' ability to replicate and divide. Eflornithine, often in combination with nifurtimox,

is mainly used as a back up for failed melarsoprol treatments in late stage HAT (Kennedy 2004). The side effects of eflornithine are similar to nifurtimox such as gastrointestinal problems as well as anemia and convulsions (Burri & Brun, 2003). However, eflornithine is expensive and treatment requires high intravenous dosages. For these reasons it is difficult to distribute and administer in rural Africa (Croft et al., 2005).

A)



B)



C)

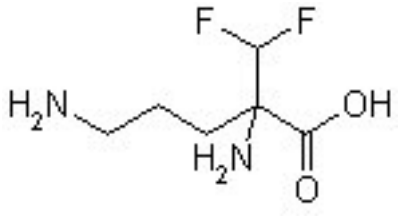


Figure 3 Molecular Structures of Melarsoprol, Nifurtimox and Eflornithine A) Melarsoprol, an arsenical derivative B) Nifurtimox, 5-nitrofuran C) Eflornithine, DMFO

Images courtesy of Laboratoire de Parasitologie Faculté de Pharmacie Lille: http://pharion.univ-lille2.fr/labos/parasito/Internat/medicam/tryp_me.html#4

Clearly, there is a desperate need for new drug technologies. One of the most current chemotherapy alternatives is administering these drugs in combinations (Priotto et al., 2006). A study done by Gerardo Priotto and colleagues for late stage West African HAT used various combinations of the afore-mentioned drugs. This study tentatively showed the best efficacy from eflornithine combined with nifurtimox. Their trial was ended abruptly, however, due to ethical reasons of increased death rate in patients undergoing treatment with some of the other drug combinations. Nifurtimox combined with melarsoprol showed a high mortality rate (Priotto et al, 2006).

Other drugs are available for treatment of HAT. Pentamidine, 1,5-bis (4-amidi-phenoxypentane) has been available for over 30 years and is normally used for treatment of first stage *gambiense* HAT (Nok 2003). It behaves as a trypanocidal compund, similar to melarsoprol except that it cannot cross the BBB, lysing the cells that uptake the molecule. Suramin was first given to sleeping sickness patients in 1922. It is a naphthalene derivative that is used for first stage *rhodesiense* infections (Nok 2003).

Problems exist with the current drugs in addition to their physical side effects. There is an increasing rate of resistance developing in HAT patients to melarsoprol (Priotto et al., 2006). A recent study on a new clinical test to identify melarsoprol-resistant trypanosome strains showed it to be a useful diagnostic tool. Resistance to the drug is due to the lack of a plasma membrane aminopurine transporter, called P2 (Stewart et al., 2005). This transporter is responsible for

internalizing melarsoprol as well as pentamidine (Nok 2003). The assay for resistance uses a fluorescent molecule, 2,5-bis-(amidinophneyl)-3,4-dimethylfuran, that is taken into the cell via the P2 transporter and fluoresces in the presence of DNA (Stewart et al., 2005). Hence the trypanosome fluoresces if it is not resistant. The paper presented by Stewart et al. showed this assay to be clinically sensitive and yield results quickly (Stewart et al., 2005).

Trypanosomiasis is a global problem adversely affecting hundreds of thousands of people. Current medical therapies are toxic, outdated, expensive and difficult to produce and distribute. The overall improvement of disease management and treatment relies on the continuation of trypanosome research. Heightened understanding of trypanosome genetics and physiology and their interaction with mammalian immune systems will provide better targets for drug therapies.

Trypanosome Life Cycle, Cellular Structure and Processes

The African trypanosome has a complex life cycle and an uncommon cellular composition and physiology that make it recalcitrant to the host immune system and drug treatment (Morgan et al., 2001). These parasitic protozoans take on various morphologies and changes in physiological characteristics throughout their developmental stages within the fly and human host (Gull 2002).

The life cycle of the African trypanosome resembles other digenetic parasites (Bastin et al., 2000). Throughout this cycle the cells take on characteristic forms that adapt to the diversity of host environments that they

encounter (Matthews, 2005). The schematic of the life cycle of the African trypanosome, depicting the various cell forms, is shown in Figure 4.

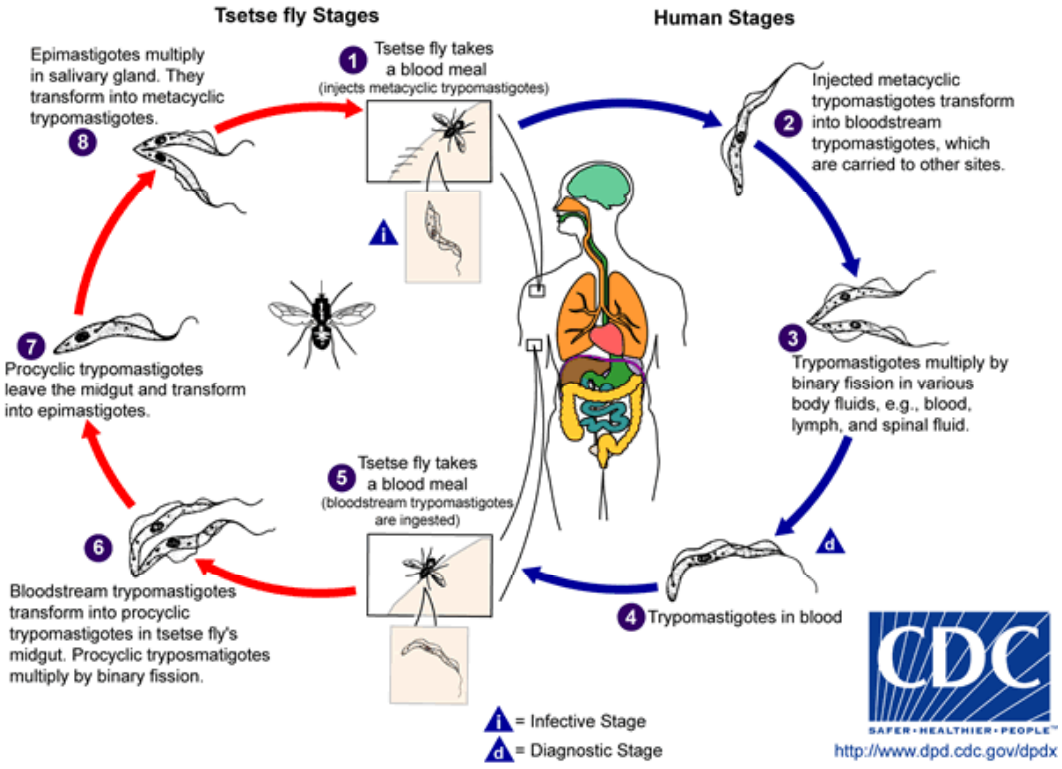


Figure 4 Life Cycle of the African Trypanosome. The various cell morphologies, proliferative and non-proliferative, are shown that occur in both fly and human hosts. Figure courtesy of Centers for Disease Control and Prevention: http://www.dpd.cdc.gov/DPDx/HTML/ImageLibrary/TrypanosomiasisAfrican_il.asp?body=S-Z/TrypanosomiasisAfrican/body_TrypanosomiasisAfrican_il5.htm

The cycle begins when a tsetse fly bites an infected person, injecting the metacyclic trypanosomes that have matured in the flebotomine (fly) salivary glands into the human host. These cells will develop into the long slender proliferative stage (number 3 in Figure 4) in the blood stream and other bodily fluids. Eventually some of these long slender trypomastigotes become short

stumpy trypomastigotes, the non-proliferative stage (number 4). It is the short stumpy stage that establishes a population in the flebotomine gut upon the next fly bite (number 5). The stumpy bloodstream form survives fly ingestion and migrates to the gut where they develop into a dividing procyclic form (number 6). Eventually they cease dividing and travel into the salivary glands. The procyclic cells become epimastigotes (number 7) in these glands and attach themselves to develop into the metacyclic form (number 8) that can be released into a new host (number 2) via the fly bite (Bastin, 2000).

Trypanosomes have an overall long and cylindrical shape that is established by a cytoskeleton comprised of polarized microtubules. Microtubules are long, polymers that are mostly made of αβ-tubulin heterodimer monomers (Kohl & Gull 1998). Figure 5 shows a trypanosome cell body with organelles. All microtubules in the trypanosome cell are arranged minus to plus end, being arranged in a helical pattern going from anterior to posterior (McKean, 2003). The trypanosome flagellum is tethered to the entire length of the cell and runs posterior to anterior (Gull, 2002).

The organelles of each cell are concentrated between the middle of the cell body and its posterior end. At the posterior end is the flagellar pocket, an opening where the flagellum exits the interior of the cell. This pocket serves as the exclusive site of endo- and exo-cytosis (Morgan et al., 2001).

Some of the trypanosome organelles move location within the cell in a tightly regulated manner that corresponds to the cell's stage of development. This

is most noticeable in the migration of the kinetoplast, located at the base of the flagellar pocket. It is found in the posterior end of bloodstream cells and further towards the anterior end in procyclics (Hendriks et al., 2000). This migration is so tightly controlled that DAPI staining of the mitochondrial DNA in the kinetoplast of cells can show its position and can be used to determine their precise stage of development (Matthews, 2005).

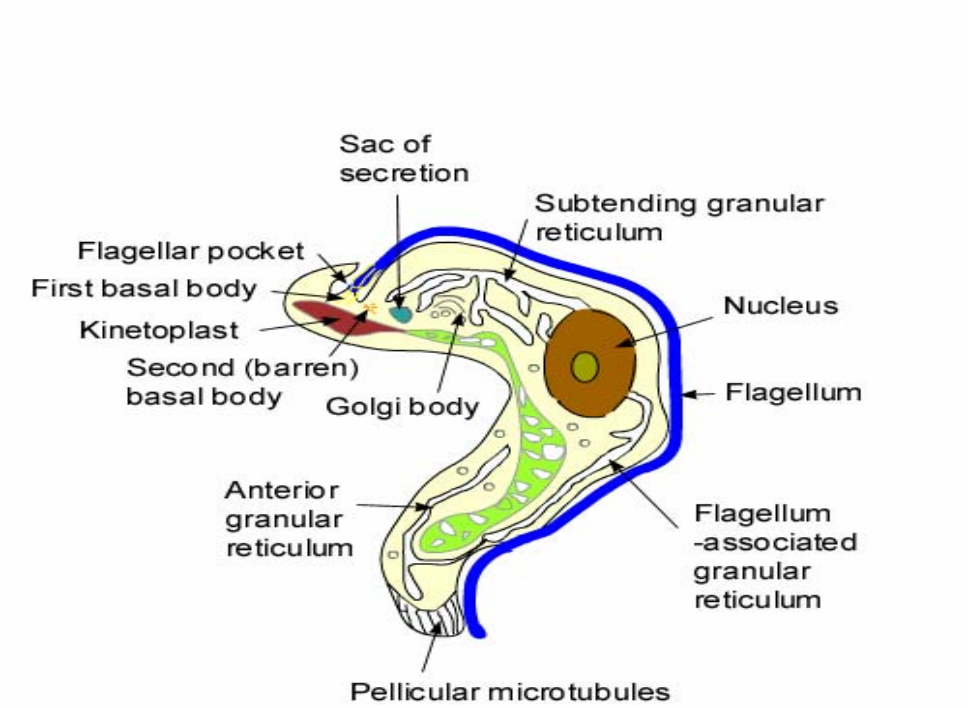


Figure 5 Shows trypanosome organelles. The cell contains a nucleus as well as a kinetoplast, both containing genetic material. The flagellar pocket, the site of all endocytosis, is near the kinetoplast. Trypanosomes have one long mitochondria, shown in green. Image courtesy of University of South Carolina School of Medicine: www.pathmicro.med.sc.edu/parasitology/blood-proto.htm

Trypanosomes have a single nucleus and mitochondrion. They also have a kinetoplast, a unique matrix of catenated mitochondrial DNA, termed kDNA (Morris et al., 2001). This matrix is an intertwined network of small and large circles of kDNA, termed mini- and maxi-circles. Trypanosome kDNA has its own replication system and contains an unusual genetic code that must be edited to produce functional mRNA (Klingbeil & Englund, 2004). Possession of the kinetoplast puts trypanosomes in the order of Kinetoplastida. All organisms of this order contain mitochondrial kinetoplasts and are largely free-living flagellated parasites.

The nucleus of *Trypanosoma brucei*, the common laboratory species used to study African sleeping sickness, contains 11 large chromosomes and more than 100 small (~50 kb) chromosomes (Matthews, 2005). These minichromosomes contain genes encoding the expansive collection of variant surface glycoproteins (VSG), the molecular technique used by the *T. brucei* to cloak itself from the mammalian immune system. Some of these genes are found in the larger chromosomes as well (Rudenko, 2000). VSG are membrane anchored, sugar-protein complexes that serve as surface molecules and are used as a highly efficient strategy for evading the host immune system (Donelson, 2003). Mammalian blood stream trypanosomes continually switch expression of these surface markers to cloak itself from immune system attack. This means that all but one of over a 1000 VSG genes are transcriptionally silent at any given time. These genes have different loci throughout many of the chromosomes, yet it

seems that only the VSG genes located at the telomeres are expressed (Donelson, 2003). It is unknown as to why this is. The high rate of change of amino acid variability between sets of VSG molecules is the molecular basis of their host evasion technique (Gull, 2002). Despite the chemical differences between each VSG type they all produce similar rod-like structures that pack tightly together, effectively covering the cell body (Donelson, 2003). Trypanosomes repeatedly switch their VSG type, meaning they are periodically changing the chemical composition of their entire outer surface (Matthews, 2005). In doing so they are able to remain one step ahead of the host's immune system. By constantly shuttling new VSG coats through their flagellar pockets (Matthews, 2005) trypanosomes completely change their identifying surface molecules and thus effectively cloak themselves from the mammalian host's alternative complement pathway and adaptive immune system efforts (Vincendeau & Bouteille, 2006). The trypanosome was the first organism in which antigenic variation was described (Stich et al., 2006). VSG gene arrangement and regulation as well as VSG molecular diversity and switching continue to be of tremendous scientific interest.

Trypanosome Flagellum

Trypanosomes are motile, flagellated cells possessing a single flagellum that contains the classic 9+2 axoneme found in all organisms with motile flagella (Hutchings et al., 2003). Despite the appearance of this highly conserved

structure there is much unknown about its proteomic composition and mechanical function.

Trypanosome flagella have conserved motor proteins, calcium-binding proteins, microtubule arrangements and other connecting and regulatory complexes found to be conserved in other flagellated organisms (Baron et al., 2006). This classic flagellar/ciliary structure is comprised of two central microtubule pairs connected to 9 outer doublet pairs via linking proteins called radial spokes (Yang et al., 2004). Flagellar wave propagation is thought to arise by the movement of dynein motors. These protein motors are attached to the central pair and walk along the outer doublets, bending the entire flagellum (Yang et al., 2004). The radial spokes are known to function in the coordination of dynein movement in *T. brucei* (Ralston et al., 2006).

Due to the possession of the classic motile axoneme, trypanosomes are used as a model system for eukaryotic flagellar research (Hutchings et al., 2002). Other single celled organisms such as *Chlamydomonas* are also commonly used in flagellar studies (Luck, 1984). The information gathered from these model organisms have far reaching implications as this conservation of structure is seen in the motile flagellum of mammals such as humans. Flagellar defects are the root cause of many human diseases (Mitchell 2005). Due to the high degree of conservation of this cellular structure between organisms it is relevant and useful to conduct flagellar research on protozoans such as *Chlamydomonas* and

Trypanosoma, organisms that are more efficiently and quickly grown in culture and are less of a moral burden when experimented upon.

Trypanosomes also possess some unique qualities about their flagella. The flagellum of *T. brucei* winds around the cell body (McKean, 2003). The attachment along the cell facilitates the unusual movement employed by trypanosomes, the entire cell spirals in a corkscrew motion to move (Hutchings et al., 2002). Indeed, the word trypanosome itself expresses this characteristic as "trypanon" means auger in Greek (Hill 2003). The wave propagation is initiated from posterior base to anterior tip, which is reversed in most other flagellated organisms (Ralston et al., 2006). Aside from motility, the trypanosome flagellum serves other physiological purposes, making it a multifunctional organ functioning in processes such as cytokinesis and host-parasite interactions (Baron et al., 2006). The area where the flagellum exits the cell body, the flagellar pocket, is the sole site of endocytosis. The flagellum is also the area where cell recognition and adhesion reactions occur (Wu et al., 1994). Trypanosomes use their flagella to attach to the epithelial cells in the tsetse fly salivary gland. It also serves as a sensory organ, responding to environmental stimuli (Ralston et al., 2006).

The paraflagellar rod (PFR) is an additional structure found in *Trypanosoma* spp. flagella. It is not present in other trypanosomatids such as *Leishmania* spp., which have free flagella (Bastin, 2000). The PFR begins where the flagellum exits the flagellar pocket and parallels the axoneme. It is a protein-

dense structure that serves as an area of attachment for the axoneme and has a half-moon shape in cross-section (Hill, 2003). The PFR is enclosed with the flagellum in the flagellar membrane and has been shown to be essential for proper motility (Gull, 2002). Intact flagellar function has been proven to be required for cell viability in general, as throughout its development the trypanosome must migrate within a given host to survive (Bastin, 2000).

In studying flagellar structure and function in trypanosomes we cannot only elucidate solutions to human diseases based upon flagellar defects but also demystify the parasite itself. As described above, HAT is a terrible disease that has not changed much in the last few decades. Its prevalence is owed to a multitude of factors but is predominantly due to a lack of safe, affordable and obtainable drug therapies. New approaches for HAT treatments are required. Understanding the molecular mechanisms involved in navigation of motility and basic flagellar motor function will help determine novel sites to target for new chemotherapies that are desperately needed to combat trypanosomiasis.

T. brucei Components of Motile Flagella (TbCMF)

A recent study from the Kent Hill research lab of UCLA sought to contrast the trypanosome genome against other organisms. This team of researchers performed a bioinformatics comparison using a genomic database to compare the genome of the trypanosome against all other flagellated organisms, specifically those with motile flagella (those possessing the 9+2 axoneme). In this search they looked for genes in *T. brucei* that pertain to conserved motile flagellar genes

found in other motile-flagellated organisms as well. Their search yielded a set of 50 proteins contained in the trypanosome genome that are conserved and unique among organisms with only motile flagella and not organisms possessing non-motile flagella. The set of corresponding genes were termed *T. brucei* Components of Motile Flagella (TbCMF). Of these 30 were found to be previously undescribed. They then performed RNA interference (RNAi, discussed below) on 41 of these genes (the 30 novel genes plus several others) to characterize their respective protein function. Almost all genes assayed showed some degree of motility defect when knocked down. The Hill team ranked the motility defects they saw into classes, numbered 1 to 4. They described Class 1 as being largely unaffected, Class 2 as mildly affected, Class 3 as moderately affected and Class 4 as severely affected. They then categorized all 41 genes into the class that best fit the phenotype displayed by each one. The study went on to examine more closely some of the more severely defective phenotypes, including using a protein database to derive hypothetical protein identities for the knockdown mutants of interest (Baron et al., 2006).

In collaboration with the Hill UCLA research group, I obtained a list of the 41 TbCMF genes they investigated in their study. This list gave a summary of each gene: its TbCMF number, Gene DB identification number, its nucleotide length, its protein's molecular weight, amino acid length and what protein it is most homologous to. It also listed the class descriptor for each. This descriptor categorizes the degree of motility defect to which the knock down of the

respective gene causes (Baron et al., 2006). I was given the opportunity to characterize a gene in the Class 3 category. The Hill team has called this TbCMF gene TbCMF 63, its Gene DB number is Tb11.01.1210. The TbCMF 63 protein is listed as being homologous to calcyphosine. Calcyphosine is a calcium-binding protein (CaBP) in the calmodulin super family, an EF-hand protein (discussed below). Calcyphosine was first isolated and characterized from dog thyroid (Lecocq et al., 1995).

In this study, I sought to characterize TbCMF 63 by elucidating its CaBP domain. Based upon my interpretation of this domain's tertiary structure several different amino acid changes were made to target residues that are hypothetically involved, directly or indirectly, in the chelation of calcium. By perturbing the protein function to produce a mutant phenotype we hope to gain insight about the role of CMF 63 in the flagellum and motility. Understanding of motile flagellar function has extreme significance to human diseases stemming from ciliary defects. Through complete knowledge of this structure and its function better treatments can be developed to combat such ailments.

EF-Hand Calcium-Binding Proteins

The family of EF-hand CaBP is one of the most common CaBP groups among eukaryotic organisms (Nelson et al., 2001). These proteins facilitate vital processes within most eukaryotic cells in cooperation with calcium ions (Nelson & Chazin, 1998). EF-hand proteins have a distinctive tertiary form (Grabarek,

2006). Yet, despite their common structural similairties these proteins form an amazing diversity of classes and participate in a variety of cellular mechanisms.

EF-hand CaBP have a conserved tertiary structure that is present throughout the large variety of classes that comprise EF-hand. The EF-hand motif is a helix-loop-helix structure that occurs in pairs (Grabarek, 2006). Each loop chelates a single calcium ion (Nelson & Chazin, 1998). Loops can vary slightly in numbers of amino acid residues that comprise them. Generally, calcium binding loops in EF-hand proteins have either 12 (canonical loop) or 14 residues (pseudo loop). These CaBP are distinguished by having different combinations of loop residues and number of loop pairs. The smallest EF-hand CaBP is D_{9k} and has only one EF-hand pair. Troponin C and calmodulin have two pairs (Nelson & Chazin, 1998). The number of pairs continues to increase among the EF-hand varieties. The S100 class of EF-hand proteins form dimers (Santamaria-Kisiel et al., 2006). This wide variation on a theme is testimony to the diversity of functions performed by EF-hand CaBP (Nelson & Chazin, 1998). The EF-hand pairs are usually cooperative, using hydrogen bonds and/or salt bridges that occur between the two loops' residues to stabilize and aide each other in calcium chelation.

Calcium ions are one of the most common intracellular signaling molecules (Grabarek, 2006). They act as second messengers in signaling pathways in neurons and function in cellular mechanisms such as gene expression, cell division and differentiation, transportation of intracellular

molecules and cell motility (Hoeflich & Ikura, 2002). Changes in intracellular concentrations of free calcium ions are the basic control mechanism behind this myriad of cellular processes (Israelson et al, 2005). EF-hand CaBP are major regulators of intracellular calcium concentration (Johnson, 2006). These proteins regulate the level of calcium by acting as buffers or modulators (responsible for calcium homeostasis), transducers or sensors in cascade events and as transporters (Johnson, 2006). All EF-hand CaBP perform this variety of functions by binding free calcium ions.

EF-hand proteins are typically involved in modulating calcium concentrations or transducing a calcium signal downstream in a biochemical event (Nelson et al., 2001). When an EF-hand CaBP transducer binds calcium it usually undergoes a conformational change. This structural alteration causes the EF-hand protein to act upon a target peptide in the reaction sequence, exposing reactive residues to act upon the target (Johnson, 2006). Calmodulin is a known EF-hand transducer.

In the eukaryotic flagellum, calcium binding events are thought to play a role in the control of wave formation in flagellar motility (Luck, 1984). A proteomic study done by Pazour et al. analyzed the eukaryotic motile flagella using *Chlamydomonas reinhardtii* as a model. From their screen, they found 27 proteins to contain EF-hand motifs (Pazour et al., 2005). Another flagellar study done by Yang et al. found calmodulin to be part of the radial spoke protein complex (Yang et al., 2004). Radial spokes have been implicated in controlling

the shape of the flagellar bend in the motile 9+2 axoneme (Ralston et al., 2006). There is also a known EF-hand CaBP flagellar gene family in *T. brucei.* The authors called this family calflagin and characterized some of the protein members within it. They hypothesize, due to their sequestration in flagella, that calflagin participates in motility (Wu et al. 1994). These studies suggest that calcium binding is an essential mechanism for ciliary and flagellar bending and that EF-hand proteins participate in this mechanism (Yang et al., 2003).

Based on this information, I believe TbCMF 63 is important to proper cell motility and hypothesize that motility defects will be seen in mutations to this gene that cause perturbations to the protein's normal function. In this study I performed site specific mutagenesis to the C-terminal end of CMF 63's putative EF-hand protein, making three separate mutations based upon my hypothesized structure of this portion of the protein. These mutations will be expressed in African trypanosomes using RNA interference techniques.

RNA Interference and Site Specific Mutagenesis

RNA interference is a newly defined powerful tool that allows for the silencing of a gene without mutations to create knock outs, when the gene of interest is physically removed or rendered untranslatable. This naturally occurring cellular mechanism exists in many eukaryotes, working to degrade dsRNA (Aigner, 2006). This process is employed for different roles in different organisms. RNAi has been shown to silence parasitic genes in plants and transposons in worms (Kong et al., 2007). The potential applications for the

exploitation of this biological mechanism extend beyond gene characterization studies (Gheysen & Vanholme, 2007). Transgenic *Arabidopsis thaliana* and tobacco have been engineered to produce dsRNA specific to hormones secreted by root-infecting nematodes. This hormone is thought to aid the infectious nature of the nematodes. These transgenic plants effectively transfer dsRNA to the worms, silencing the production of this hormone. They have subsequently shown a greater resistance to nematode infection as compared to non-transgenic plants (Gheysen & Vanholme, 2007). The potential for mammalian drug therapies based on RNAi is clearly recognized. Researchers are striving to better understand the specific nature of RNAi in different organisms and identify the best delivery methods for it. Clinical applications of RNAi for human disease treatments give tremendous promise for future drug therapies and drug delivery systems (Kong et al., 2007).

The general mechanism of RNAi involves the breakdown and incorporation of double stranded RNA (dsRNA) into a protein complex that can then target and destroy other existing RNA with the same sequence (Kong et al., 2007). An overview of this mechanism is shown in Figure 6. The RNAi mechanism is possessed by many eukaryotic cells and is thought to exist as a defense against viral or transposon invasion (Montgomery, 2006). The dsRNA is recognized by an ATP driven enzyme called Dicer. This enzyme processes the dsRNA into small interference RNA (siRNA) (Aigner, 2006). These siRNA are then incorporated into a protein called the RNA Induced Silencing Complex

(RISC) (Kong et al, 2007). With the help of RNA helicase the RISC protein unwinds the siRNA into single strands that remain in RISC and go on to bind with messenger RNA (mRNA) within the cell. The binding of RISC to mRNA fragments causes their cleavage and subsequent degradation (Kong et al., 2007).

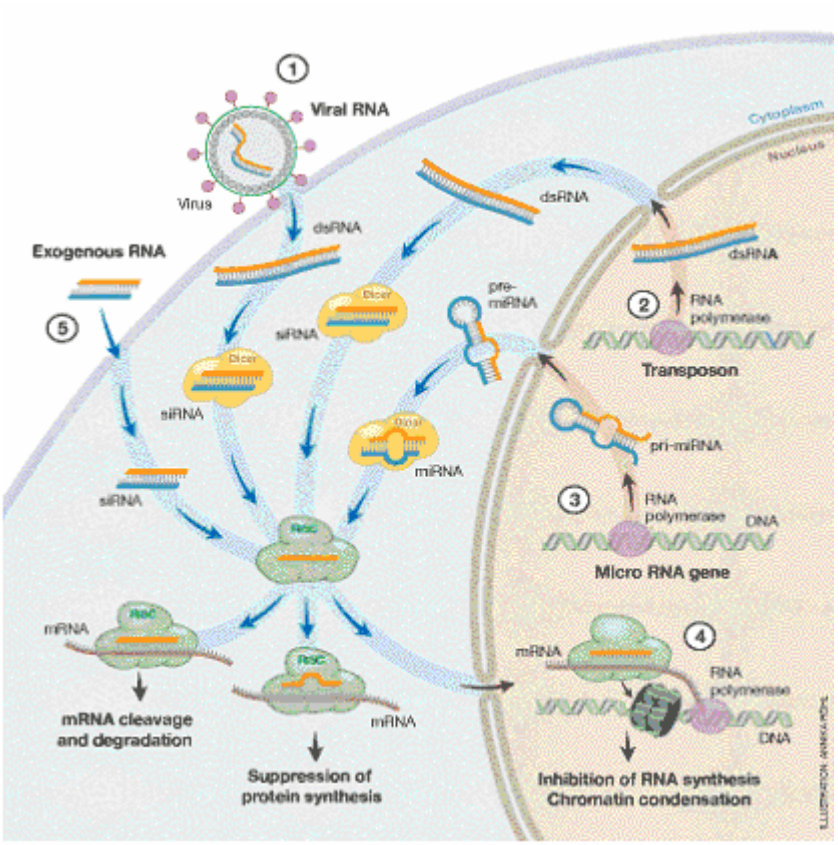


Figure 6 Overview of RNAi Pathway and dsRNA Delivery Options
http://nobelprize.org/nobel_prizes/medicine/laureates/2006/adv.html

The dsRNA can be introduced *in vitro* into a given cell in a variety of ways. Direct injection of dsRNA or siRNA is often employed as well as the introduction of exogenous DNA designed to produce the desired dsRNA fragments (Montgomery, 2006).

Experimental Approach

RNAi Mutatgenesis Scheme Overview

While RNA interference is usually spoken of as completely knocking down the expression of a gene it can also be used in a more sophisticated manner. In this study we are using RNAi to aid in the expression of my set of EF-hand mutations. To the recombinant *T. brucei* strain 29-13 (Wirtz et al., 1999) we will introduce two vectors, one to knock down the expression of endogenous CMF 63 and the other to carry the mutated version of the gene whose expression will be induced.

The Details of Our RNAi Technique

The knockdown of the endogenous CMF 63 and the expression of each mutated version will be done by transfecting these plasmids into a recombinant strain of *T. brucei*. This strain, 29-13, is based on the wild-type *T. brucei* 427 strain. It was engineered to have multiple drug reporter genes. The 29-13 strain was created by cloning two plasmids containing promoters, T7 RNA polymerase (T7RNAP) and a tetracycline repressor (TetR), into the α-β *TUBULIN* locus of the genomic DNA in the 427 strain. The TetR is an area of DNA that binds binds an transscriptional inhibiting protein that releases upon biding with tertracycline. Each of these plasmids carried a drug resistance gene to behave as a reporter, hygromycin and G418 respectively. A promoter is a regulatory region of DNA that binds RNA polymerase, the enzyme that serves to express the gene

downstream of the given promoter. Hence, the 29-13 strain is grown on hygromycin and G418 to maintain the pure culture strain.

A plasmid containing the endogenous 3'UTR of CMF 63 will be introduced into the *T. brucei* 29-13 strain. This vector contains opposing T7 promoters. The endogenous 3' UTR is inserted in between these promoters, which in turn will allow for expression of this region. The excessive copies of mRNA of the endogenous 3' UTR will effectively silence the transcription of the entire gene. This plasmid contains a phleomycin resistance gene as well. Thus, upon transfection of the 3' UTR plasmid into 29-13 trypanosomes they will require a third drug to be grown in culture.

Lastly, each mutated gene will be introduced into *T. brucei* via another plasmid that will be incorporated into the chromosomal DNA via its trypanosomal rRNA spacer. Also, each mutation will be placed in front of a heterologous 3' UTR so as not to be targeted by the RNAi knockdown of the endogenous 3'UTR. This plasmid contains a resistance to the drug puromycin, requiring a fourth drug for culture. Using the TetR feature of the 29-13 strain makes it possible for all mutations to be tetracycline drug induced and have the ability to be switched on and off.

Transfections, Inductions and Screening for Mutants

All transfections, inductions and mutant screenings will be carried out at UCLA in the Kent Hill lab. They are designing the plasmid containing the 3'

UTR of endogenous CMF 63. I will send my mutated versions of CMF 63 ligated in the puromycin vector to UCLA.

Methods and Materials

Cloning CMF 63

The target ORF of CMF 63 was amplified from purified *T. brucei brucei,* strain 29-13, gDNA (obtained from Kent Hill, UCLA). Primers, each 29 bases in length to include restriction site and spacer end, were designed to amplify the entire ORF, 1,659 bp in length. Primers were ordered through IDT (www.idtdna.com). See Table 1 below.

Table 1 Primers for PCR Amplification of CMF 63

CMF 63 (*Xba* I) primer	CMF 63 (*Hin* dIII) primer
5'CACTCTAGATTACATATTAATTCTAATGT3'	5'CACAAGCTTATGAACGAATTAAAGAACCC3'

Amplification was performed using 1µl 5U/µl GoTaq DNA polymerase and 5 µl 5x buffer from Promega (www.promega.com), 1 µl of each of the 5 µM primers, 2 µl 25 µM dNTP, 2 µl 30ng/ml 29-13 gDNA template and 2 µl of 25 mM $MgCl_2$brought up to 25 µl volume reaction. All PCR reactions were carried out in a Techne Thermocycler in the following sequence:

- denature at 94° C for 1 minute
- then 30 times each:
  - denature at 94°C for 30 seconds
  - anneal at 40° C for 30 seconds
  - 72° C for 2 minutes
- final extension at 72° C for 10 minutes
- hold at 4° C

PCR was performed with a low annealing temperature of 40°C due to unequal base pair ratios between the primers. Magnesium chloride was required for the reaction in order to facilitate primer binding. PCR was analyzed by gel agarose electrophoresis and positive products were extracted from the gel using QIAquick gel extraction kit (Qiagen, www1.qiagen.com) and a concentration was determined using Nanodrop nd-1000 spectrophotometer (NanoDrop, www.nanodrop.com). The plasmid vector, pKH10, was used for ligation with amplified CMF 63. A figure of pKH10 is given in Appendix A. This plasmid was provided by Kent Hill and contains a GFP gene flanked by *Xba*I and *Hin*dIII sites. CMF 63 and pKH10 were digested with restriction enzymes *Xba*I and *Hin*dIII (Promega) and incubated at 37°C overnight.

The reactions were heat inactivated to kill the enzymes. The pKH10 plasmid was then prepared for ligation with the CMF 63 insert. A phosphatase reaction was then performed on pKH10 in order to cleave the phosphate groups on the 5' ends of the strands to prevent the plasmids from religating. The plasmid was incubated at 37°C for one hour with the enzyme alkaline phosphatase and then stopped with EDTA. The reaction was either extracted with phenol-chloroform and ethanol precipitated or run on a gel and extracted with QIAqiuck (Qiagen) to purify the DNA away from the alkaline phosphatase. Concentration was determined using a Nanodrop nd-1000 spectrophotometer.

Ligation of pKH10 and CMF 63 was performed using T4 DNA ligase from New England Biolabs (www.neb.com). Two insert to vector ratios were

used, 6:1 and 10:1. Reactions were carried out for 15 minutes at room temperature. The cloned plasmids were transformed into $CaCl_2$ competent DH5α *E. coli* cells and plated on Luria broth, a nutrient rich medium, containing ampicillin. Colonies were then screened for transformants.

Purification of pET160-CMF63

A new vector was obtained from the Kent Hill (UCLA). This plasmid contained CMF 63 cloned into a plasmid called pET160/GW/D-TOPO (Invitrogen, www.invitrogen.com). We referred to this sample of plasmid as pET160-CMF63. A figure of this plasmid is given in Appendix B. A large-scale DNA preparation was performed using the Promega Wizard DNA extraction kit and the concentration of the purified pET160-CMF63 was determined using Nanodrop nd-1000 spectrophotometer.

Determination of the C-terminal EF-Hand Domain of TbCMF 63

The entire amino acid sequence for CMF 63 was compared to homologous proteins using the BLASTp algorithm (Altschul et al., 1990) on the BLAST internet site (http://130.14.29.110/BLAST). As mentioned above, the C-terminus is less variable in amino acid content than the N-terminus. Thus, the C-terminal end of CMF 63 CaBP domain was studied to determine where the functional EF-hand motifs were located. The length of CMF 63 CaBP domain was elucidated from the C- terminus. From this, determination of the functional EF-hand domains of the C-terminus of CMF 63 was made and three mutations were designed.

Design Primers for Site-Directed Mutations

To perform site specific mutagenesis primers for each mutation, shown in Table 1, were made following the protocol of the QuikChange Site-Directed Mutagenesis kit (Stratagene, www.stratagene.com).

Table 2 Primers for CMF 63 EF-hand Mutations with Corresponding Endogenous DNA Sequence

Mutation Name	Endogenous DNA Sequence	Mutated DNA Primer Sequences
D477A	5'gcaatcgagagc**gac**gaacgtttcttc3'	5'gcaatcgagagc**gcg**gaacgtttcttc cgttagctctcgcgccttgcaaagaag5'
E433A	5'cccgatgtagag**gag**ggcattgccagt3'	5'cccgatgtagag**gcg**ggcattgccagt gggctacatctccgcccgtaacggtca5'
S456W	5'atgaacagtacc**tcg**tctgtaacaactgat3'	5'atgaacagtacc**gcg**tctgtaacaactgat tacttgtcatggcgcagacattgttgacta5'

The base pairs indicated in **bold** indicate the area of change. In the Endogenous DNA Sequence column the red base pairs show the original code and those id the Mutated DNA Primer Sequences show the changes made to produce mutations.

The primers for mutations D477A and E433A were 27 base pairs in length with the changed codon in the center. The third primer made for S456W was 30 base pairs in length, bearing the changed codon near the middle as well.

QuikChange Site-Directed Mutagenesis

All mutagenesis reactions were performed via the protocol given in QuikChange Site-Directed Mutagenesis Kit. All reactions were prepared as follows:

- 5 µl of 10x reaction buffer
- 48.25 ng of dsDNA template pET160-CMF63

- 125 ng of oligonucleotide primer #1
- 125 ng of oligonucleotide primer #2
- 1 µl of dNTP mix
- brought all reactions up to 50 µl total volume with dH_20

All reactions ran in a Techne Thermocycler following these cycle parameters:

- denature at 95°C for 30 seconds
- then 12 times each:
  - denature at 95° C for 30 seconds
  - anneal at 55°C for 1 minute
  - 68° C for 7.5 minutes
- hold at 4° C

All reaction samples were then digested with 1 µl of *Dpn* I restriction enzyme and transformed into supercompetent $CaCl_2$ XL-1 Blue *E. coli* cells (Stratagene). To verify that mutations were performed successfully DNA was purified and a sample of each was sent for sequencing and was also analyzed by digestion with restriction enzymes. *Eco*RV was used for the digest as it cuts twice in pET160-CMF63, producing two bands: one band 3261 bp long and the other 4237 bp. After each mutation is confirmed, they will be ligated into a cloning vector for *T. brucei* genes that contains a puromycin resistance cassette provided by Kent Hill.

Results

PCR of CMF 63

Figure 7 shows the results for PCR amplification of CMF 63. As seen in the gel, the product runs just above 1.6 kb, corresponding with the proper length of endogenous CMF 63 which is 1659 bp long.

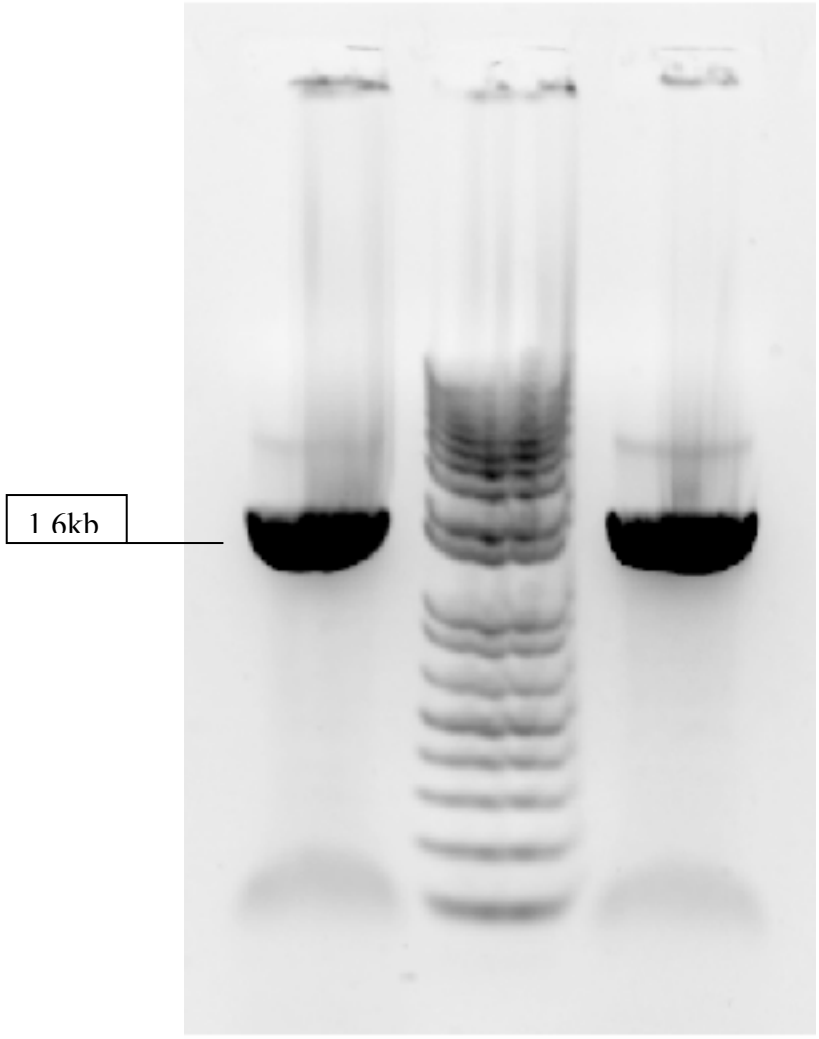


Figure 7 PCR Product of CMF 63 As seen in the above gel a positive product was obtained for the PCR amplification reaction of endogenous CMF 63. The gene is 1659 bp in length. Due to the large amount of product obtained it is difficult to see that the product bands run just above the 1.65 kb marker band. A 12 kb ladder was used for this analysis.

C-Terminal EF Hand Domain of TbCMF 63

The entire CMF 63 protein is 552 amino acids in length. BLASTp results showed the beginning (C-terminal end) of the entire EF-hand domain to start around valine at position 340 (V340) the outer most α-helix beginning at tryptophan (W) 489. For a complete table of amino acids see Appendix C. The results of the database query demonstrated the putative protein's N-terminus varied significantly between homologues. Residue V340 is most frequently given by BLASTp results to be the N-terminal starting amino acid of the first calcium binding loop flanking α-helix. The sequence for dog thyroid calcyphosine was used as the primary template for determining homologous protein domains because it is well characterized and has classic canonical EF-hand motifs. The BLASTp results are shown below in Figure 8A. The canine homologue showed 29% amino acid identity over the 143 amino acids in the comparison. Also, a distance tree, a cladogram showing the degree of similarity between CMF 63 and the database homologues, is given for the BLASTp results and is shown in Figure 8B. The unnamed protein is tbCMF63. On the tree it has been matched with its GeneDB number, Tb 11.01.1210. The tree shows CMF 63 to have close identity with other trypanosomes. The tree was compiled by using a maximum sequence difference of 0.75.

A)

calcyphosine [Canis familiaris]
sp|P10463|CAYP1_CANFA Gene info Calcyphosin (Calcyphosine) (Thyroid protein p24) (TPP) (Protein 5)

Score = 77.4 bits (189), Expect = 2e-12, Method: Composition-based stats.
Identities = 42/143 (29%), Positives = 78/143 (54%), Gaps = 2/143 (1%)

```
Query  348  GISLMGVTIHPGELD V IF KKL  DRVGNGFVVAQEFLRELRCELP QSRLQGVI SAFQQLVIE  407
            G++ +G+   +     E  +   + ++     DR  G+G +      +EFLR  LR   +  Q+R      + +AF   +L
Sbjct  49   GLAELGLVLDTAEAEGVCRRWDRD GSGTLD L EEFLRALRPPMSQAREAVIAAAFAKLDRS  108

Query  408  GGGSVDYKDMLNLFVFNACFHPDVEEG IASREE II F D FINCWPNMNSTSSVTTDMFVAYY  467
            G  G V       D+        V++       HP  V+  G    + EE ++     F++  +  +            VT       F    YY
Sbjct  109  GDG  VVTVDDLRG--VYSGRTHPKVQSGEWTEEEVLRRFLDNFDSSEKDGQVTLAEFQDYY  166

Query  468  TDVSPAIES DE R F FK M LKRCWKI  490
            +  VS  ++++ DE    F    M +       W++
Sbjct  167  SGVSASMDTDEEFVA MMTSAWQL  189
```

B)

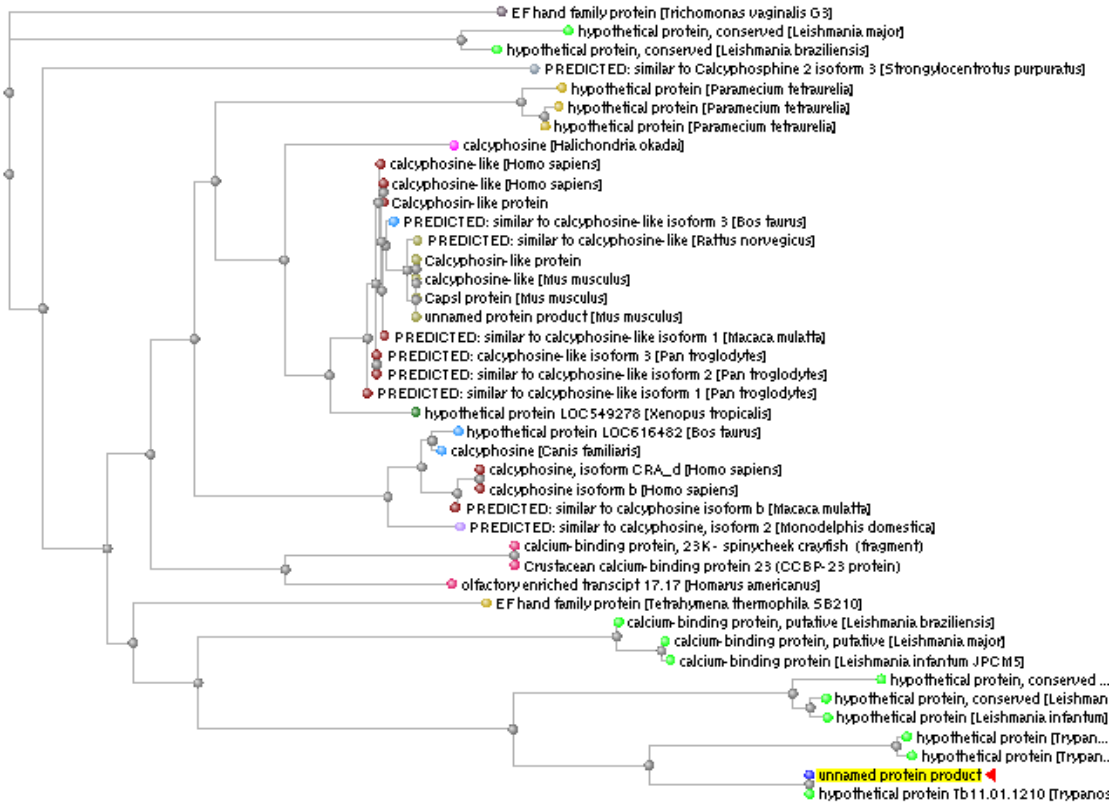


Figure 8 BLASTp Results A) Shows BLAST results for *Canis familiaris* calcyphosine (subject) aligned with the end of the sequence of CMF 63 (query) (http://130.14.29.110/BLAST/Blast.cgi#50979090) B) A distance tree produced from the results of the BLASTp algorithm. Color code: CMF 63 kinetoplastids ciliates mammals sponges crustaceans eukaryotes rodents amphibians primates marsupials echinoderms (http://www.ncbi.nlm.nih.gov/blast/treeview/blast_tree_view.)

Calcyphosine was first isolated and identified in dog thyroid and contains classic EF-hand amino acid structure (Lecocq et al., 1995). A study by Lecocq et al.

found calcyphosine to be a calmodulin-like EF-hand CaBP that is a phosphorylated substrate in the cyclic AMP cascade in dog thyroid cells (Lecocq et al., 1995). Comparing the dog calcyphosine reference to CMF 63 I was able to establish the loop and helical regions in the primary sequence of CMF 63. The tertiary structure of this hypothetical protein was hence elucidated and is shown in Figure 9 below.

Like calmodulin (CaM), CMF 63 was shown to have two canonical (12 residues each) binding loops at the C-terminal end. Two red ovals have been drawn into each calcium binding loop to indicate the position of the chelated calcium ion. The inner flanking vertical strands are 7 amino acid long α-helices and are connected by the linker loop, 8 residues in length. The left outer-most vertical strand represents approximately half the central α-helix that connects the two EF-hand domains, the C- to N-terminus. The right outer-most vertical strand is the C-terminal flanking α-helix. The numbered residues are the targets for mutagenesis.

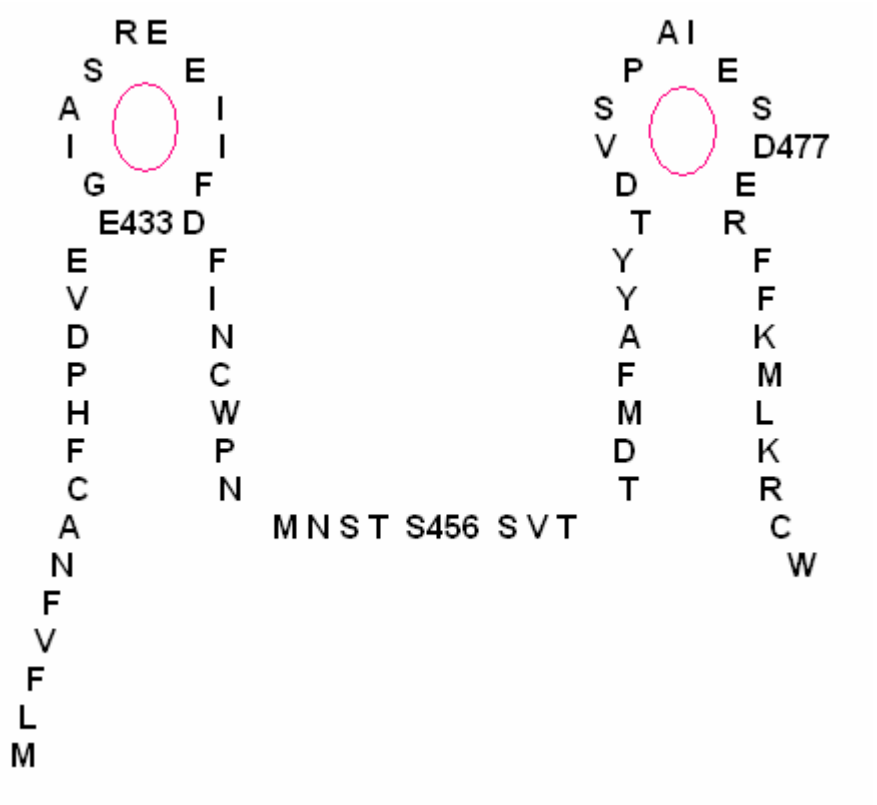


Figure 9 C-terminal End of CMF 63 In this figure, the vertical amino acid strands represent α-helices and the horizontal strand the linker loop between the two calcium binding loops. The red circles indicate the position of individual calcium cations as they are chelated in the binding loops. The numbered amino acids indicate the target residues for mutagenesis.

Design of Mutants

The first mutation was designed to perturb the hydrogen bonding between the C-terminal calcium binding loops. Changing residue aspartic acid 477 (D477) to an alanine (A) will disrupt the salt-bridge normally formed between this residue and arginine 438. This mutation was termed D477A. The second mutation was made to weaken the calcium chelation in one of the two loops, changing residue glutamic acid 433 (E433) to an alanine. The carboxylate side chain of E433 is directly involved in binding calcium in the loop. This mutation was termed

E433A. The third mutation was created to give a silent phenotype. An amino acid change in the linker portion between the two EF-hand domains was selected, changing residue serine 456 (S456) to a tryptophan (W). The linker is an area that is highly variable in all EF-hand proteins and may be functionally tolerant to residue changes. This mutation was termed S456W.

Discussion

The Trouble with pKH10

Initially, I experienced a very low yield in the PCR amplification of CMF 63. This problem was overcome by lowering the annealing temperature to 40°C and adding magnesium. The low temperature was required due to the difference in purine:pyrimidine ratios between the two primers. Addition of magnesium helps coordinate the incorporation of dNTPs in the PCR reaction solution. Magnesium ions are chelated by dNTPs. The positive charges on the magnesium attract the magnesium:dNTP complex to the negatively charged phosphate groups of the denatured DNA strands. This attraction serves to create more efficient amplification reaction as the dNTPs are more readily added on.

Despite the success of the PCR amplification of CMF 63 it would not transform into competent *E. coli* cells after ligation with pKH10. To overcome problems with recovery of DNA after phenol-chloroform extraction, I ran digested, phosphatase-treated pKH10 in a gel directly after the phosphatase reaction in order to purify the plasmid away from the enzyme. Gel extraction of the plasmid after this step produced a better yield than the chloroform extraction. However I was still able to obtain only a few viable transformants after using this plasmid preparation for ligation with CMF63. Fortunately, I was able to obtain a new plasmid from Kent Hill, pET160-CMF63, that already included CMF 63. This vector had already been successfully transformed into *E. coli* at UCLA. Due

to the availability of the vector I did not further pursue the problems I was having using pKH10.

Validity of the EF-hand Model

BLAST results revealed EF-hand homology towards the end of CMF 63, having V340 as the most frequent N-terminus. The strongest identities seen in the 103 proteins from the database aligned with CMF 63 query pertained to EF-hand calcium binding loops. These proteins gave a percent identity ranging from 56% within other trypanosomes to percentages in the mid-twenties as compared to the most distant organisms (see Figure 8B). The strongest regions of identity within all subject organisms corresponded to the EF-hand calcium binding loops. As discussed in the introduction, the calcium binding loops are generally the most highly conserved portion of EF-hand proteins. The identity seen in these areas of CMF 63 supports the proposed EF-hand model. The length of CMF 63, as determined by the BLASTp algorithm, is about 150 amino acids. This finding also supports the accuracy of the interpretation that CMF 63 is a EF-hand CaBP as 150 amino acids is the average length for calmodulin.

Significance of the Three Mutations and Their Predicted Phenotypes

In these three mutations I sought to create structural changes to the CaBP domain of CMF 63. However, I expect for only 2 of the 3 mutations to produce defects in the motility of *T. brucei*. These predicted motility defects may manifest in different ways depending on the affected flagellar component. Mutant phenotypes produced by previous flagellar protein studies have shown that the

type of defect experienced by the cell depends on the disabled protein(s). An RNAi knockdown study was done on trypanin, a protein found in the flagellum of *T. brucei* (Hutchings et al., 2002). In this study, ablation of trypanin expression showed flagellar detachment from the cell. The flagellar detachment in turn caused the cells to tumble without direction although they maintained normal wave propagations (Hutchings et al., 2002). In another study, RNAi was used to knockdown some of the protein components of the axonemal radial spokes and some of those in the microtubule central pair in *T. brucei* (Ralston et al., 2006). These component proteins included trypanin. This study determined trypanin to be part of the dynein regulatory complex (DRC), a complex that communicates with the central microtubule pair to generate normal flagellar beat. Ablations of trypanin, another DRC component and a component in the central microtubule pair produced severe motility defects. Trypanosomes lacking intact radial spokes and microtubule central pairs were described as immotile, with acutely disrupted wave propagation. Cytokinesis was disrupted as well in these mutations (Ralston et al., 2006). The target of my mutations is calmodulin-like CMF 63. EF-hand CaBP have been shown to be present and function in the flagellum of *T. brucei* (Wu et al., 1994). By disturbing the calcium binding function of CMF 63 I believe I will disrupt signaling for wave propagation in the flagella. Calmodulin is known for participating in transduction pathways and may play such a role in flagellar beat. Hence, erratic flagellar beating is a likely mutant phenotype to be seen in this study. I also do not expect a severe phenotype in my mutations as I

am not completely silencing the expression of the gene. CMF 63 will still be present in the mutants and will most likely continue to function in all three mutations. Hence, with partial function remaining, the mutant phenotype should not be as severe as if I had completely knocked down the expression of the gene.

The first mutation, D477A, was designed to interfere with the stability of the tertiary structure of the two calcium binding loops. As previously stated, these loops classically form hydrogen bonds between the residues in the loops that are not responsible for calcium chelation. The stability of the tertiary structure has been shown to play a role not only in the control of the protein's conformational changes but also in the act of calcium binding itself. Hence, this mutation has the potential to affect the two main functional aspects of CMF 63 and might yield the most severe motility defect phenotype of the three mutations generated.

The second mutation is expected to interfere with the calcium chelation of one of the calcium binding loops. E433A creates the loss of a calcium ligating bond within this loop. This residue has no other interaction with any other residue in the rest of the EF-hand domain. As this will only partially perturb one of four calcium binding regions within this protein I predict the possible phenotype to be moderate in loss of proper control of motility.

For the third mutation I made an amino acid change in the linking loop between the two calcium binding loops of the C-terminal EF-hand domain. As stated above, this region is often the most variable in amino acid content between

given EF-hand proteins. Hypothetically, whatever change may be introduced into this sequence should have little to no effect on the CaBP overall function. By hypothesizing that the S456W mutant will have no net effect on the native protein's operation I am testing two things: firstly, the accuracy of my interpretation of the C-terminus of CMF 63's EF-hand domain and secondly, the validity of the assertion that the linker loop is in fact highly variable. Hence, I introduced a radically different amino acid, tryptophan, from the original serine in the middle of this loop. Due to the severity of the amino acid change, however, it is entirely possible a mutant phenotype could be seen. Tryptophan is a hydrophobic residue that will exert some hydrophobic force. This force may be sufficient to cause structural changes to the linker loop as it tries to bury itself away from the other solvated residues. This in turn may disrupt the conformations of the attached calcium binding loops. By imposing this severe change I hope to characterize the limitations on the structural requirements for this area of the protein.

Previous studies have explored CaM point mutations *in vivo* and found them to produce defective phenotypes. A 2004 study produced a lethal phenotype in pupal *Drosophila*. Valine 91 was changed to a glycine (V91G) in the sole endogenous CaM gene found in *Drosophila* (Wang et al., 2004). The authors could not distinguish any differences concerning affinity for calcium binding nor overall secondary and tertiary structure between wild-type CaM and that of V91G. However, a mild destabilization of the C-terminal end was seen in regard

to V91G interacting with it's target peptide during muscle contraction. The authors concluded that lethality of the phenotype was due to deregulation of calcium fluctuations that led to muscle hypercontraction and subsequent failure (Wang et al., 2004).

Recall the African trypanosome flagellum functions in cellular processes other than locomotion. The trypanosome flagellum serves as an attachment site for the cells that migrate to the fly salivary gland. This migration is necessary for maturation and transfer to a new host (Bastin et al., 2000). It is possible, considering the various physiological changes that a single trypanosome undergoes, for a mutation to give a phenotype in some forms and not others. Hence, it is possible that any of the above described mutations could present a phenotype in the procyclic form. This was shown in a study using RNA interference to knock down several different flagellar proteins (Broadhead et al., 2006). These mutations showed proper flagellar function was required for viability in the bloodstream form. Ablation of these genes disrupted cytokinesis, producing convoluted, multi-flagellated monster cells that eventually died. This phenotype occurred only in bloodstream form of *T. brucei* and not procyclic (Broadhead et al., 2006). Another study of flagellar protein knock downs in *T. brucei* produced similar defects (Ralston, et al. 2006). The base of the flagellum, the flagellar pocket, is also the site of all vesicular transport (Matthews, 2005). This is another mechanistic role performed by the flagellum that could be subject

to regulatory disruption. Due to the trypanosome flagella having more than one function defects other than motility are possible.

Identity of N-Terminal Portion of CMF 63

The CaBP domain of CMF 63 only comprises approximately one third of the entire protein. Complete elucidation of CMF 63's EF-hand function will lie in determining the identity of the remaining two thirds of the protein domain. I entered this portion of the protein, 324 amino acids in length, into a separate BLASTp search in attempt to characterize it. The identities resulting from the search are two receptors, an inositol 1,4,5-trisphosphate (IP_3) and ryanodine receptor. The search yielded a percent identity ranging from 53% - 21% among the 24 homologies. The domain is specifically identified as the ligand-binding region in IP_3 receptor and the N-terminal in the ryanodine receptor (Marchler-Bauer & Bryant, 2004). IP_3 and ryanodine receptors are two families of calcium ion mediating channels (Nelson & Cox, 2005). These channels are usually located in the ER membrane, the major area of calcium sequestration, allowing calcium to flow into the cytosol upon stimulation (Nakayama et al., 2004). They are often linked to the activation of neurotransmitter-gated receptors and voltage-gated calcium channels in the plasma membrane. This coupling of intra- and extracellular calcium signals allows for molecular communication within a cell and between cells. The determination of different neuronal signals are known to occur through this pathway (Marchler-Bauer & Bryant, 2004). Many of these channels are stimulated by cleaved molecules called secondary messengers that

arise from extracellular binding events (Nelson & Cox, 2005). Different hormones and neurotransmitters that bind various cell surface receptors activate the plasma membrane bound enzyme phospholipase C . This enzyme, when activated, cleaves the plasma membrane lipid phosphatidylinositol 4,5-bisphosphate into two secondary messengers, diacylglycerol and IP_3. IP_3 diffuses through the cytosol to the ER where it binds to a group of IP_3 receptors on calcium channels and stimulates calcium release (Nelson &Cox, 2005). Calmodulin (CaM) has been shown to work with this system as well. A study was done on transient receptor potential (Trp) channels, the calcium channels in the *Drosophila* ER, looking specifically at their IP_3 receptors (Tang et al., 2001). In the Trp channels studied, IP_3 and CaM were found to share binding sites on all receptors. It is thought that IP_3 competes with CaM for these binding sites. CaM has an inhibitory effect on the Trp channel. A functional study was performed on one of the channels, determining that CaM is actually tethered to the receptor. The authors believe permanent positioning of CaM at the channel receptor serves to prevent accidental channel stimulation and calcium release (Tang et al., 2001).

For trypanosome flagella, this is somewhat mysterious. There is no known ER component within the flagellum, no place for ER calcium channels to reside. The Tang paper discusses the activation of SOC (store-operated channels) (Tang et al., 2001). These channels are responsible for the influx of calcium from the extracellular space. The SOC channels are very similar in structure and regulation to the ER channels. While little is know about either type of channel,

the paper suggests that both systems work together using CaM and IP_3 to regulate calcium intake and release (Tang et al., 2001). It is possible that there are unknown IP_3-like channels residing within the flagellar membrane or elsewhere in the flagellar region. It is understood that while flagella/cilia lack core cell organelles within the flagella membrane compartment the membrane itself possesses common receptor-ligand interactions (Wang et al., 2006) as well as transporter proteins (Stewart et al., 2005). It is not a stretch of the imagination that the flagellar membrane may retain some common ion channels as well as receptors.

The Importance of CaM and Its Role in Flagellar Function

Clearly, CaM plays an important role in modulating intracellular calcium, an important cellular signal, as shown by the number of studies presented above. The reoccurrence of its conserved structure among various eukaryotes demonstrates it has been an evolutionarily important protein. It was also shown to directly participate in motility in the Baron et al. study done by the Kent Hill lab (Baron et al., 2006). Many of the other studies given as examples in this paper support the wide range of important functions performed by CaM. This includes direct participation in cell motility through calcium ion signaling. In order to understand the calcium mechanism of cell motility it is necessary to identify each protein that participates in it. Collectively, this shows the importance of confirming the identification of CMF 63 and characterizing its function within the flagella.

Provided more time for this project I would like to further probe the function of CMF 63 by designing further mutations. To begin, I believe mutations carried out in the N-terminus of the CaM domain would be informational. Comparing the phenotypes of C- versus N-terminal mutations could indicate a greater functional importance of one end than the other or show them to be equally important. I would also like to compare the site-specific EF-hand mutations to knock downs of the entire EF-hand domain. I am curious to know what the difference in phenotype is between a defective and missing CMF 63 EF-hand. In turn, knockdowns of just the IP_3 CMF 63 domain may provide useful information as to function when compared to observations of the latter. Combining this data with the localization information produced by the GFP and antibody tagging planned for CMF 63 by Kent Hill should provide a better picture for the role of this gene in motility.

Elucidating the details of trypanosome motility has the potential to help further the understanding of a myriad of human ciliary-based diseases. Motile cilia are found throughout the human body (ovaries, sperm, kidneys, lining of the respiratory tract and the brain). Diseases such as juvenile myoclonic epilepsy, infertility (Baron et al., 2006) and polycystic kidney disease (Morgan et al., 2005) are just a few examples of diseases based on ciliary defects. These diseases plus others make this study and others like it contemporarily relevant and important.

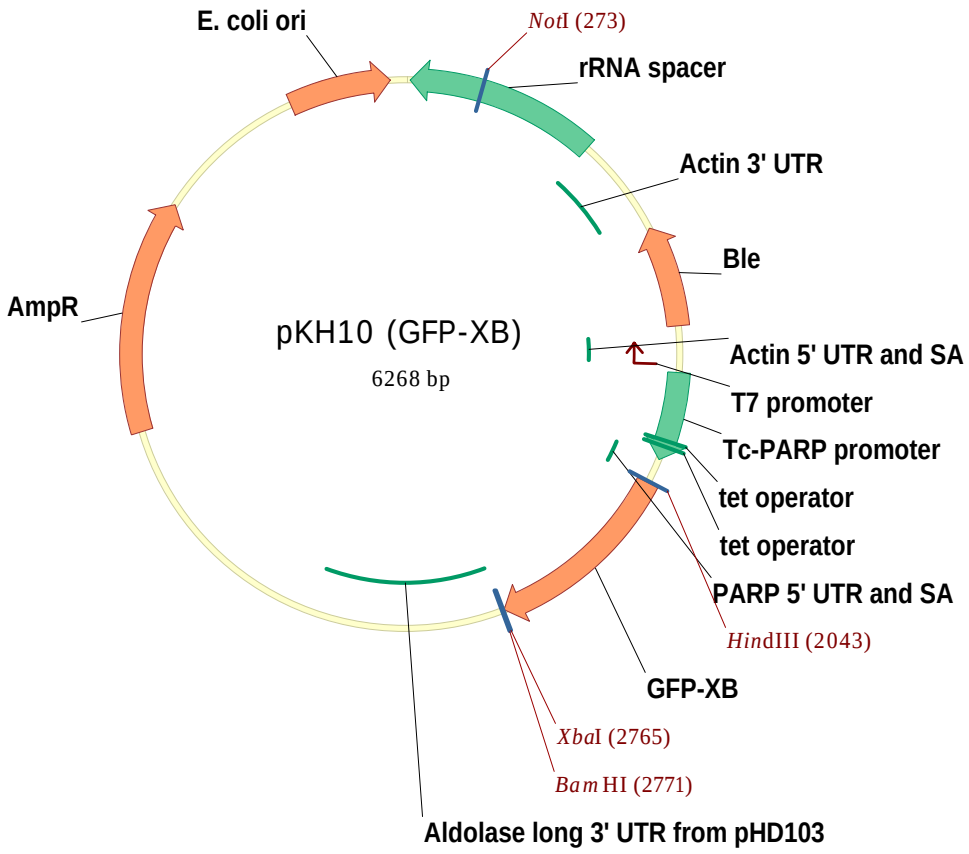
A more complete understanding of the cell biology of the trypanosome is necessary to further efforts to manage and treat trypanosomiasis. At the moment

there is not enough known about the organism to provide for better therapies. The lack of any recent developments for new drugs have left the doctors and researchers who deal with trypanosomiasis to experiment with mixtures of the old toxic ones (Priotto et al., 2006). The VSG system of the trypanosome make it nearly impossible to design an effective trypanosomiasis vaccine (Donelson, 2003). The day is approaching, however, where new treatments will be an absolute necessity as more and more infected people show immunity to the current ones.

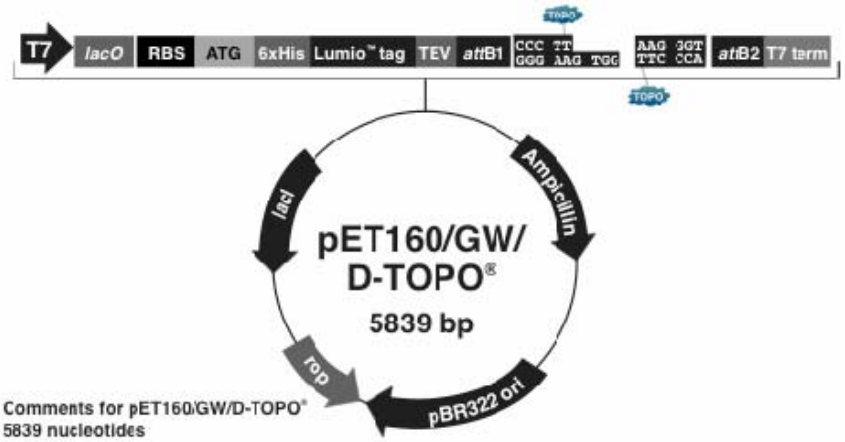
Approaches to managing incidence of infection can be improved through trypanosome research as well. There is a need to better prevent disease in human and cattle population. Devising strategies to control the parasite while in the fly or other animal reservoirs is a key component to managing and reducing human infections. Tactics for such management are developed through careful research of the organism of interest.

APPENDIX

Appendix A. pKH10 plasmid



Appendix B. pET160-CMF63 vector

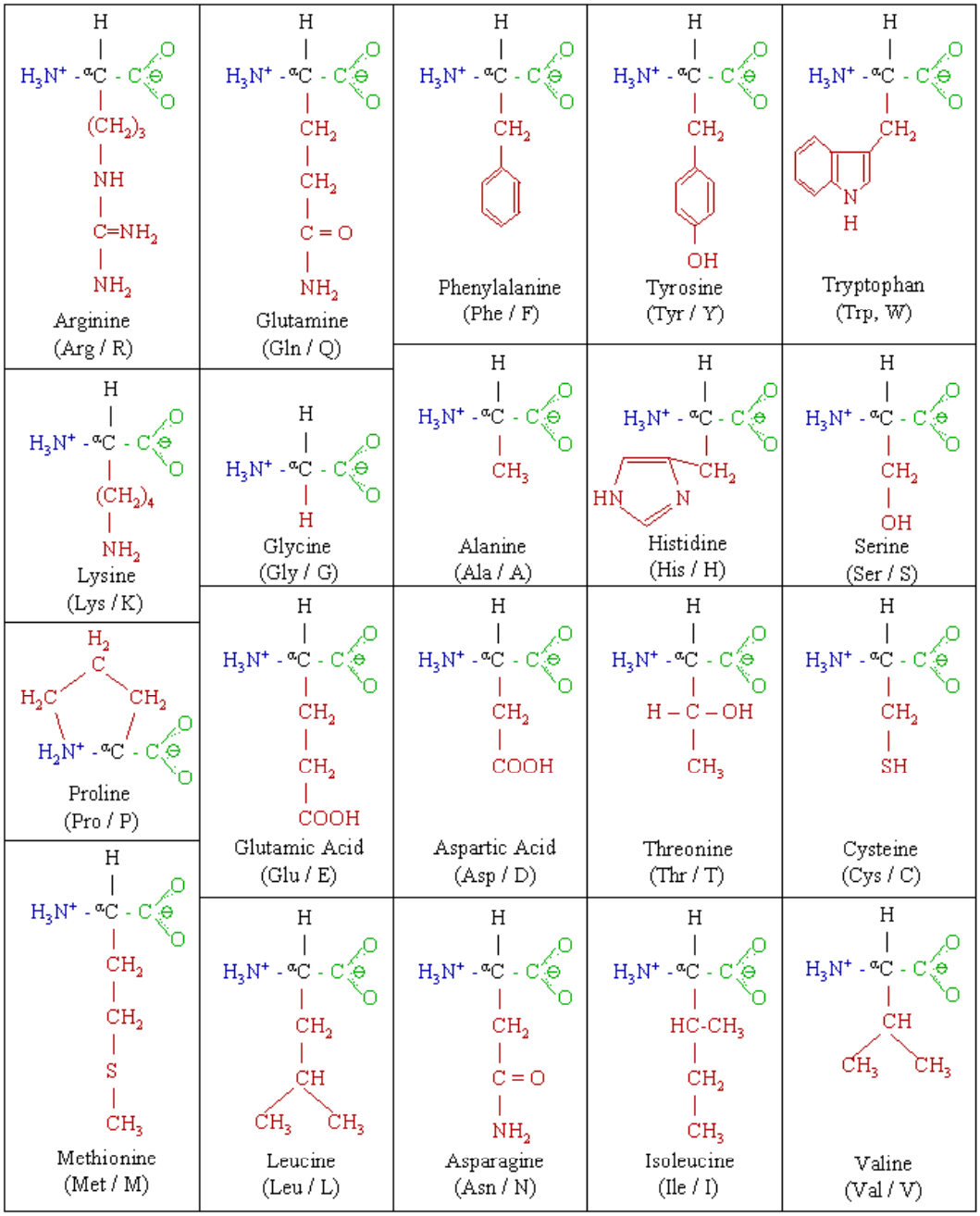


Comments for pET160/GW/D-TOPO®
5839 nucleotides

T7 promoter/priming site: bases 21-40
lac operator (lacO): bases 40-64
Ribosome binding site (RBS): bases 94-101
Initiation ATG: bases 109-111
Polyhistidine (6x-his) region: bases 112-129
Lumio™ tag: bases 142-159
TEV recognition site: bases 169-189
*att*B1 site: bases 196-220
TOPO® recognition site1: bases 232-236
Overhang sequence (c): bases 237-240
TOPO® recognition site 2: bases 241-245
*att*B2 site: bases 257-281
T7 transcription termination region: bases 307-435
T7 reverse priming site: bases 346-365
bla promoter: bases 740-838
Ampicillin (*bla*) resistance gene: bases 839-1699
pBR322 origin: bases 1844-2517
POP ORF (c): bases 2888-3079
lacI ORF (c): bases 4391-5482

(c) = complementary strand

Appendix C. Table of Amino Acids



http://www.biocrawler.com/encyclopedia/Aminoacids

REFERENCES

Aigner, A. (2006). Delivery systems for the direct application of siRNAs to induce RNA interference (RNAi) in vivo. Journal of biomedicine & biotechnology, 2006(4), 71659.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology, 215(3), 403-410.

Baron, D. M., Ralston, K. S., Kabututu, Z. P., & Hill, K. L. (2007). Functional genomics in *Trypanosoma brucei* identifies evolutionarily conserved components of motile flagella. Journal of cell science, 120(Pt 3), 478-491.

Bastin, P., Pullen, T. J., Moreira-Leite, F. F., & Gull, K. (2000). Inside and outside of the trypanosome flagellum:A multifunctional organelle. Microbes and infection / Institut Pasteur, 2(15), 1865-1874.

Bisser, S., N'Siesi, F. X., Lejon, V., Preux, P. M., Van Nieuwenhove, S., Miaka Mia Bilenge, C., et al. (2007). Equivalence trial of melarsoprol and nifurtimox monotherapy and combination therapy for the treatment of second-stage *Trypanosoma brucei gambiense* sleeping sickness. The Journal of infectious diseases, 195(3), 322-329.

Broadhead, R., Dawe, H. R., Farr, H., Griffiths, S., Hart, S. R., Portman, N., et al. (2006). Flagellar motility is required for the viability of the bloodstream trypanosome. Nature, 440(7081), 224-227.

Burri, C., & Brun, R. (2003). Eflornithine for the treatment of human African trypanosomiasis. Parasitology research, 90 Supp 1, S49-52.

Chappuis, F., Loutan, L., Simarro, P., Lejon, V., & Buscher, P. (2005). Options for field diagnosis of human African trypanosomiasis. Clinical microbiology reviews, 18(1), 133-146.

Courtioux, B., Boda, C., Vatunga, G., Pervieux, L., Josenando, T., M'Eyi, P. M., et al. (2006). A link between chemokine levels and disease severity in human African trypanosomiasis. International journal for parasitology, 36(9), 1057-1065.

Croft, S. L., Barrett, M. P., & Urbina, J. A. (2005). Chemotherapy of trypanosomiases and leishmaniasis. Trends in parasitology, 21(11), 508-512.

Deborggraeve, S., Claes, F., Laurent, T., Mertens, P., Leclipteux, T., Dujardin, J. C., et al. (2006). Molecular dipstick test for diagnosis of sleeping sickness. Journal of clinical microbiology, 44(8), 2884-2889.

Donelson, J. E. (2003). Antigenic variation and the African trypanosome genome. Acta Tropica, 85(3), 391-404.

Garcia, A., Courtin, D., Solano, P., Koffi, M., & Jamonneau, V. (2006). Human african trypanosomiasis: Connecting parasite and host genetics. Trends in parasitology, 22(9), 405-409.

Gheysen, G., & Vanholme, B. (2007). RNAi from plants to nematodes. Trends in biotechnology, 25(3), 89-92.

Grabarek, Z. (2006). Structural basis for diversity of the EF-hand calcium-binding proteins. Journal of Molecular Biology, 359(3), 509-525.

Gull, K. (2002). The cell biology of parasitism in *Trypanosoma brucei*: Insights and drug targets from genomic approaches? Current pharmaceutical design, 8(4), 241-256.

Hendriks, E., van Deursen, F. J., Wilson, J., Sarkar, M., Timms, M., & Matthews, K. R. (2000). Life-cycle differentiation in *Trypanosoma brucei*: Molecules and mutants. Biochemical Society transactions, 28(5), 531-536.

Hill, K. L. (2003). Biology and mechanism of trypanosome cell motility. Eukaryotic cell, 2(2), 200-208.

Hoeflich, K. P., & Ikura, M. (2002). Calmodulin in action: Diversity in target recognition and activation mechanisms. Cell, 108(6), 739-742.

Hutchings, N. R., Donelson, J. E., & Hill, K. L. (2002). Trypanin is a cytoskeletal linker protein and is required for cell motility in African trypanosomes. The Journal of cell biology, 156(5), 867-877.

Israelson, A., Arzoine, L., Abu-hamad, S., Khodorkovsky, V., & Shoshan-Barmatz, V. (2005). A photoactivable probe for calcium binding proteins. Chemistry & biology, 12(11), 1169-1178.

Johnson, C. K. (2006). Calmodulin, conformational states, and calcium signaling. A single-molecule perspective. Biochemistry, 45(48), 14233-14246.

Kennedy, P. G. (2004). Human African trypanosomiasis of the CNS: Current issues and challenges. The Journal of clinical investigation, 113(4), 496-504.

Klingbeil, M. M., & Englund, P. T. (2004). Closing the gaps in kinetoplast DNA network replication. Proceedings of the National Academy of Sciences of the United States of America, 101(13), 4333-4334.

Kohl, L., & Gull, K. (1998). Molecular architecture of the trypanosome cytoskeleton. Molecular and biochemical parasitology, 93(1), 1-9.

Kong, Y., Ruan, L., Ma, L., Cui, Y., Wang, J. M., & Le, Y. (2007). RNA interference as a novel and powerful tool in immunopharmacological research. International immunopharmacology, 7(4), 417-426.

Lecocq, R., Lamy, F., Erneux, C., & Dumont, J. E. (1995). Rapid purification and identification of calcyphosine, a Ca(2+)-binding protein phosphorylated by protein kinase A. The Biochemical journal, 306 (Pt 1)(Pt 1), 147-151.

Luck, D. J. (1984). Genetic and biochemical dissection of the eucaryotic flagellum. The Journal of cell biology, 98(3), 789-794.

Lutumba, P., Robays, J., Miaka mia Bilenge, C., Mesu, V. K., Molisho, D., Declercq, J., et al. (2005). Trypanosomiasis control, Democratic Republic of Congo, 1993-2003. Emerging infectious diseases, 11(9), 1382-1388.

Marchler-Bauer A, Bryant SH (2004), "CD-Search: protein domain annotations on the fly.", Nucleic Acids Res.32(W)327-331.

Matthews, K. R. (2005). The developmental cell biology of *Trypanosoma brucei*. Journal of cell science, 118(Pt 2), 283-290.

McKean, P. G. (2003). Coordination of cell cycle and cytokinesis in *Trypanosoma brucei*. Current opinion in microbiology, 6(6), 600-607.

Mitchell, D.R. (2005). Regulation of eukaryotic flagellar motility. International symposium on interdisciplinary science, 755, 130-136.

Montgomery, M. K. (2006). RNA interference: Unraveling a mystery. Nature structural & molecular biology, 13(12), 1039-1041.

Morgan, G. W., Allen, C. L., Jeffries, T. R., Hollinshead, M., & Field, M. C. (2001). Developmental and morphological regulation of clathrin-mediated endocytosis in *Trypanosoma brucei*. Journal of cell science, 114(Pt 14), 2605-2615.

Morris, J. C., Drew, M. E., Klingbeil, M. M., Motyka, S. A., Saxowsky, T. T., Wang, Z., et al. (2001). Replication of kinetoplast DNA: An update for the new millennium. International journal for parasitology, 31(5-6), 453-458.

Nakayama, T., Hattori, M., Uchida, K., Nakamura, T., Tateishi, Y., Bannai, H., et al. (2004). The regulatory domain of the inositol 1,4,5-trisphosphate receptor is necessary to keep the channel domain closed: Possible physiological significance of specific cleavage by caspase 3. The Biochemical journal, 377(Pt 2), 299-307.

Nelson, D.L. & Cox, M.M. (2005). Principles of biochemistry. W.H. Freeman and Company. New York, NY.

Nelson, M. R., & Chazin, W. J. (1998). Structures of EF-hand Ca(2+)-binding proteins: Diversity in the organization, packing and response to Ca2+ binding. Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine, 11(4), 297-318.

Nelson, M. R., Thulin, E., Fagan, P. A., Forsen, S., & Chazin, W. J. (2002). The EF-hand domain: A globally cooperative structural unit. Protein science : a publication of the Protein Society, 11(2), 198-205.

Nok, A. J. (2003). Arsenicals (melarsoprol), pentamidine and suramin in the treatment of human african trypanosomiasis. Parasitology research, 90(1), 71-79.

Pazour, G. J., Agrin, N., Leszyk, J., & Witman, G. B. (2005). Proteomic analysis of a eukaryotic cilium. The Journal of cell biology, 170(1), 103-113.

Pepin, J. (2007). Combination therapy for sleeping sickness: A wake-up call. The Journal of infectious diseases, 195(3), 311-313.

Pepin, J., Milord, F., Meurice, F., Ethier, L., Loko, L., & Mpia, B. (1992). High-dose nifurtimox for arseno-resistant *Trypanosoma brucei gambiense* sleeping sickness: An open trial in central Zaire. Transactions of the Royal Society of Tropical Medicine and Hygiene, 86(3), 254-256.

Ralston, K. S., Lerner, A. G., Diener, D. R., & Hill, K. L. (2006). Flagellar motility contributes to cytokinesis in trypanosoma brucei and is modulated by an evolutionarily conserved dynein regulatory system. Eukaryotic cell, 5(4), 696-711.

Rudenko, G. (2000). The polymorphic telomeres of the African trypanosome *Trypanosoma brucei*. Biochemical Society transactions, 28(5), 536-540.

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

Santamaria-Kisiel, L., Rintala-Dempsey, A. C., & Shaw, G. S. (2006). Calcium-dependent and -independent interactions of the S100 protein family. The Biochemical journal, 396(2), 201-214.

Stewart, M. L., Krishna, S., Burchmore, R. J., Brun, R., de Koning, H. P., Boykin, D. W., et al. (2005). Detection of arsenical drug resistance in *Trypanosoma brucei* with a simple fluorescence test. Lancet, 366(9484), 486-487.

Stich, A., Abel, P. M., & Krishna, S. (2002). Human African trypanosomiasis. BMJ (Clinical research ed.), 325(7357), 203-206.

Vincendeau, P., & Bouteille, B. (2006). Immunology and immunopathology of african trypanosomiasis. Anais da Academia Brasileira de Ciencias, 78(4), 645-665.

Wang, B., Martin, S. R., Newman, R. A., Hamilton, S. L., Shea, M. A., Bayley, P. M., et al. (2004). Biochemical properties of V91G calmodulin: A calmodulin point mutation that deregulates muscle contraction in *Drosophila*. Protein science : a publication of the Protein Society, 13(12), 3285-3297.

Wu, Y., Deford, J., Benjamin, R., Lee, M. G., & Ruben, L. (1994). The gene family of EF-hand calcium-binding proteins from the flagellum of *Trypanosoma brucei*. The Biochemical journal, 304 (Pt 3)(Pt 3), 833-841.

Yang, P., Yang, C., & Sale, W. S. (2004). Flagellar radial spoke protein 2 is a calmodulin binding protein required for motility in *Chlamydomonas reinhardti*i. Eukaryotic cell, 3(1), 72-81.