Studies on the expression of the major cell surface molecules of insect forms of *Trypanosoma congolense*, a major parasite of cattle in Africa

by

Bianca C. Loveless
B.Sc., University of Victoria, 2006

A Thesis Submitted in Partial Fulfillment of the
Requirement for the Degree of

MASTER OF SCIENCE

in the Department of Biochemistry and Microbiology

Studies on the expression of the major cell surface molecules of insect forms of *Trypanosoma congolense*, a major parasite of cattle in Africa

by

Bianca C. Loveless
B.Sc., University of Victoria, 2006

Supervisory Committee

Dr. Terry Pearson, Supervisor
(Department of Biochemistry and Microbiology)

Dr. Martin Boulanger, Department Member
(Department of Biochemistry and Microbiology)

Dr. Caroline Cameron, Department Member
(Department of Biochemistry and Microbiology)

Dr. Francis Choy, Outside Department Member
(Department of Biology)

Supervisory Committee
Dr. Terry Pearson, Supervisor
(Department of Biochemistry and Microbiology)
Dr. Martin Boulanger, Department Member
(Department of Biochemistry and Microbiology)
Dr. Caroline Cameron, Department Member
(Department of Biochemistry and Microbiology)
Dr. Francis Choy, Outside Department Member
(Department of Biology)

Abstract

African trypanosomes are protozoan parasites that cause African trypanosomiasis, diseases that affect humans and their livestock. Not only has trypanosomiasis had an overwhelming effect on the development of tropical Africa in the past, but it also constitutes one of the most significant present economic problems of the continent. Trypanosomes alternate between a mammalian host and a tsetse vector using a complex life cycle. In the mammalian host the trypanosomes live as bloodstream forms (BSFs) that are so proficient at antigenic variation, and thus host immune system evasion, that no suitable vaccine candidates have yet been identified. In contrast, the lifecycle stages that exist in the tsetse vector do not undergo antigenic variation. This potentially makes the vector-occupying trypanosomes much better targets for control if strategies can be devised to disrupt their lifecycle in the vector or to interfere with their transmission to mammalian hosts.

The primary impediment to developing strategies for disruption of trypanosome life cycles in tsetse is a lack of understanding of the molecular basis of trypanosome-tsetse interactions. Although several major surface molecules have been identified on insect form trypanosomes, these have not been well studied due to a lack of appropriate antibody

probes and to the difficulty in obtaining sufficient quantities of the different parasite life cycle stages required for such molecular studies.

My thesis research was focused on developing and using monoclonal antibody probes for analysis of expression of major surface molecules of *Trypanosoma congolense,* a serious pathogen of cattle in Africa. I used this species of trypanosome since in addition to being a socioeconomically important parasite, all four of its major life cycle stages can be grown *in vitro* in amounts sufficient for immunochemical analysis. I successfully derived and characterized monoclonal antibodies that were useful for detecting the three major surface proteins of *T. congolense* insect forms: glutamic acid/alanine rich protein (GARP), the *T. congolense* heptapeptide repeat protein (TcHRP) and congolense epimastogote specific protein (CESP). Selected monoclonal antibody probes were then employed for expression analysis of these molecules throughout the parasite life cycle using *in vitro* grown trypanosomes and parasites taken directly from infected tsetse. In addition, I determined the peptide epitopes for two of my GARP-specific monoclonal antibodies and in collaboration with Dr. Martin Boulanger and Jeremy Mason was able to localize the epitopes on a high resolution three-dimensional structure obtained by X-ray crystallography. This allowed us to derive a model that describes the orientation of GARP in the trypanosome surface membrane and explains the possible structure-function relationships involved in replacement of the bloodstream form variant surface glycoprotein (VSG) by GARP as trypanosomes differentiate in the tsetse vector after a bloodmeal.

Table of Contents

List of Figures

Chapter 1.

Chapter 2.

Chapter 3.

Chapter 4.

Chapter 5.

List of Tables

Acknowledgments

I would like to express my deepest gratitude to my supervisor Dr. Terry Pearson for his valuable advice, guidance and support of this work. Dr. Pearson has gone beyond the call of duty, and has been an inspiration to me, and those in our lab, when we have looked to him for guidance and advice in our scientific pursuits. My time in the Pearson lab will be fondly remembered and I consider myself extremely lucky to have had the opportunity to study alongside Dr. Pearson.

I also wish to express my gratitude to the official referees of my dissertation work: Dr. Caroline Cameron, Dr. Francis Choy, Dr. Martin Boulanger and Dr. Patrick von Aderkas. All of whose valuable comments and criticism were helpful in refining the draft version of the dissertation into its final form. I wish to relay further gratitude Dr. Martin Boulanger, who gave his time and expertise to extend my project beyond its original scope by crystallizing GARP.

Sincere thanks to all my friends and colleagues in: the Pearson lab, at the University of Victoria Proteomics center, and abroad, who made working on this project all the more enjoyable, and fruitful, by providing their expertise during the collection of data and laboratory work. Moreover, special thanks are due to the grad secretaries and support staff at the University of Victoria for their assistance.

Finally, I am particularly grateful to my husband, Cameron Loveless, for being my tower of strength and encouraging me in all the stages of this work. His constant and continuous support has been immeasurable on this journey. Special thanks to my mother, Delyse Tomaselli, for wisely guiding our family from South Africa to Canada, where we have been able to realize our dreams. During my studies at the University of Victoria our family has grown with the addition of our 20 month-old son, Hiram, and a baby which we are expecting one month from now. I am certain I will look back over these years that I was a student with fond memories, made more special in that it was done while growing our family.

Chapter 1. Introduction

Africa is a continent of great natural beauty and contrasts, home to vast natural resources, tropical rain forests, grassy rolling savanna, forested highlands and the largest desert in the world, the Sahara. Unfortunately, Africa is also afflicted with corrupt governments, war, poverty, malnutrition and the spread of deadly diseases. For these reasons and more, Africa remains the world's poorest and most underdeveloped continent, despite its natural wealth. According to the World Health Organization (WHO) Special Program for Research and Training in Tropical Diseases, the people of Africa suffer from all six major tropical diseases: malaria, schistosomiasis, filariasis, leprosy, Leishmaniasis and trypanosomiasis. The measurable influence of these uncontrolled scourges on Africa can be observed when infectious diseases attack crops, livestock and people, ultimately causing starvation, impaired economic development and at worst, destabilization of entire countries.

At the heart of Africa's struggle against poverty lies African trypanosomiasis. First described in the fourteenth century (see below), Human African Trypanosomiasis (HAT; also called African sleeping sickness) affects people and Animal African Trypanosomiasis (AAT) affects livestock and wild animals. These trypanosomiases are caused by salivarian protozoan parasites of the genus *Trypanosoma* that live and divide extracellularly in blood and tissue fluids of their mammalian hosts and are transmitted primarily by tsetse (*Glossina* spp.) the infamous insect vector. The prevalence and distribution of trypanosomiasis in Africa corresponds to the range of tsetse and comprises 37 sub-Saharan countries, many of which are among the poorest in the world. Eight million square kilometers of Africa are

infested with tsetse (Figure 1.1) making almost the whole of that area unproductive for animal husbandry (Fiennes, R. 1970). Much of the best-watered and most fertile land in sub-Saharan Africa is tsetse-infested.

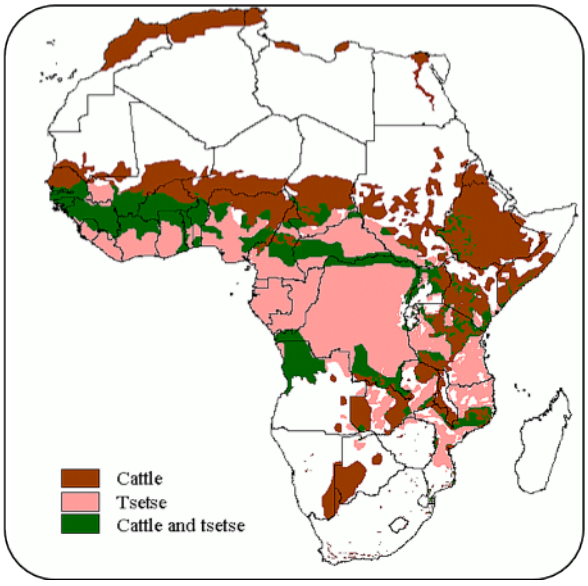

Figure 1.1. Geographic distribution of tsetse and cattle in Africa. (www.genomics.liv.ac.uk/tryps/problem.html)

Trypanosomiasis threatens human and livestock health and agricultural production, probably more so than any other single disease, thus severely repressing rural development and poverty alleviation.

1.1 History of Human African Trypanosomiasis

There are two forms of human African sleeping sickness: a chronic form prevalent in West Africa caused by *Trypanosoma brucei gambiense* and an acute form confined mainly to East and Southern Africa caused by *Trypanosoma brucei rhodesiense*. Both forms of the disease are characterized by two distinct phases: an early phase with nondescript

symptoms of nausea, lethargy and fever and a late phase, after the trypanosomes cross the blood-brain barrier, where there are disruptions in biological rhythms and sleeping patterns that result in stupor and coma, thus the name sleeping sickness. The late stage symptoms include loss of concentration and coordination, irritability, tremors, increased muscle rigidity and tonicity and behavioral changes consistent with mania or psychosis, speech disorders and seizures. All cases of human sleeping sickness are fatal if left untreated, although there are anecdotal reports of a few individuals harbouring trypanosome infections for years with no or few apparent symptoms.

The earliest recorded account of sleeping sickness came from the historical writings of Ibn Khaldoun who wrote of the death of King Diata II, Sultan of Mali in 1373, who suffered from lethargy (De Raadt, 1999). It was, however, not until 1903 when Dr. David Bruce correctly identified trypanosomes and their tsetse vectors as the causative agents of the disease. The earliest recorded major epidemics of sleeping sickness occurred in Uganda and Congo from 1896 to 1908, where roughly 500,000 people were estimated to have died in the Congo Basin and approximately 300,000 in the Busoga district of Uganda. With the Rift Valley transecting the country, Uganda is in the precarious position of having foci of both *T. b. rhodesiense* and *T. b. gambiense* forms of the disease. This resulted in two other major epidemics of sleeping sickness - one in the late 1940's and another starting in 1980. Both lasted several years. Throughout West Africa, smaller epidemics of sleeping sickness rapidly spread from Senegal to Cameroon during the 1920's, then died down by the late 1940's. By the mid 1960s, after extensive control efforts by the colonial powers, the disease had almost disappeared from many parts of Africa. With this early success came relaxed surveillance. This, coupled with the rise in political instability after independence of many

of the affected countries, saw the re-emergence of the disease in numerous areas over the last thirty years, with several endemic foci remaining in eastern and northwestern Uganda, Tanzania and elsewhere (WHO Media centre, 2006). According to the WHO, in 1995 there were an estimated 55 million people at risk of sleeping sickness in Africa, with 300,000 to 500,000 new cases each year (Smith *et al*., 1998; Hide, 1999; Welburn *et al*., 2001). By 2005, surveillance had been reinforced and the number of new cases reported throughout the continent had substantially reduced. Due to both an increase in interest and funding for neglected diseases from organizations like the Bill and Melinda Gates Foundation, the estimated number of cases is currently between 50,000 and 70,000 *per annum*, a significant improvement over the previous decade (WHO Media centre, 2006).

1.2 Animal African Trypanosomiasis

Animal trypanosomiasis, caused by several trypanosome species and carried with higher prevalence and by a greater number of *Glossina* species is, surprisingly, of much greater consequence agriculturally and economically than human trypanosomiasis. The disease in cattle, also known as nagana, results in severe losses in the productivity of domestic livestock (which are highly susceptible) due to poor growth, weight loss, low milk yield, reduced capacity for work, infertility, adult mortality, calf mortality and subsequent depressed herd growth (Mattioli and Slingenbergh, 2008). The subsequent impairment of the development of animal agriculture and sustainable food systems results in malnutrition throughout many areas of tropical Africa. Nagana has an estimated annual economic impact of US $4.5 billion to the African economy due to losses in milk, meat and wool yields (Kristjanson *et al.,* 1999). Generally speaking, trypanosome infections that threaten

livestock are 100 to 150-fold higher in *G. morsitans* than the trypanosome infections that cause human trypanosomiasis (Jordan, 1976).

Trypanosomes pathogenic to livestock include: *T. (Duttonella) vivax* (Haag *et al.,* 1998)*, T. brucei brucei* which is unable to infect humans like its companion *Trypanozoon* species *T. b. rhodesiense* and *T. b. gambiense* (Oli *et al.,* 2006) and three species belonging to the subgenus *Nannomonas: T. congolense* (Broden, 1904), *T. simiae* (Bruce *et al.,* 1912) and *T. godfreyi* (McNamara *et al.,* 1994). *T. congolense* is the most economically important due to its broad host range and wide geographical distribution.

Nagana is taken from the Zulu language and is an apt word meaning "poorly". This chronic wasting disease of livestock is primary the result of infection with *Trypanosoma congolense*. It is this species in particular that threatens upwards of 40 million cattle and kills 3 million each year. Despite its significance, *T. congolense* remains relatively unstudied when compared to the *T. brucei* group of parasites. My thesis research is focused on *T. congolense* the most important killer of cattle in sub-Saharan Africa.

1.3 Tsetse

Tsetse are large biting flies that are found only in Africa and are the biological vectors of trypanosomes (Figure 1.2). Current classifications place all 23 species and 8 subspecies of tsetse thus far identified in a single genus named *Glossina.* The genus is divided into three distinct clades: *morsitans, palpalis* and *fusca.* The *morsitans* group is the most economically important as they preferentially feed on livestock and wildlife found in the open savanna grasslands. The name tsetse (pronounced tsee-tsee) is derived from the

noise that the fly makes during flight and means, "fly" in the Tswana language. To avoid redundancy, throughout this thesis I will refer to them as tsetse rather than tsetse fly.

Each tsetse clade is classified according to preferential habitats: riverine (*Palpalis* group*)*, savannah-woodlands (*Morsitans* group*)* and forest (*Fusca* group*)*. Only the *palpalis* and *morsitans* groups are considered to be medically important because they are vectors of *T. b. brucei* spp*.* The *fusca* group is the most primitive clade and contains no species that are vectors of trypanosomiasis. Tsetse territory, which covers a third of Africa south of the Sahara (~9 million km^2) (Budd, 1999), has precluded much of the best-watered and most fertile land from cultivation and productive husbandry. As a consequence, millions of people are condemned to the futility of cultivating poor soils while fertile land lies fallow.

Figure 1.2. Image of tsetse.
Note the upward position of the proboscis. When tsetse take a blood meal, the proboscis is lowered and used to pierce the skin of the mammalian host. The blood is obtained through the hypopharynx (slender red tube seen inserted through the skin, the red colour is from the erythrocytes of the host).(tsetse image from www.dfidahp.org.uk).

Tsetse are unlike many other insects, which invest security in numbers by producing huge amounts of offspring. Instead, female tsetse are ovoviviparous, giving birth to 6-12

mature larvae in a normal lifespan of 4 months (Hoffman, 1954). The female mates in the first days of life, seldom more than once, and stores sperm in pockets in her abdomen that she releases each time she ovulates. In a normal life span she produces on average three to six offspring. Male tsetse, which are very promiscuous, have a much reduced life span of only about four weeks. Also, unlike most insects that break down their food for sugar, the tsetse's metabolism requires protein for nutrition. Proportionate to its body size it has to ingest a massive amount of blood in order to meet its food needs. In fact, tsetse are ranked the most licentious hematophagous insects in the world with respect to their lack of host specificity. The fly becomes the host for trypanosomes after feeding on the blood of an infected animal, including humans. The period from the first ingestion of trypanosome-infected blood to the appearance of the mammal-infective metacyclic forms is 12 to 21 days. The inoculum must contain a minimum of 300 to 450 individual trypanosomes for infection to be successful, and may contain up to 40,000 parasites (Hoare, 1970). Less than 90 percent of tsetse become carriers of the parasites (Welburn and Maudlin, 1999; Gibson and Bailey, 2003). However, once infective metatrypanosomes (metacyclic forms) are present, the fly remains infective.

In the last century, the mass slaughter of game, destruction of bush and spraying of DDT were strategies in an unsuccessful war against tsetse. Today, efforts are centered on reducing the tsetse population by insect trapping in combination with the sterile insect technique. Tsetse trapping involves hundreds of insecticide-impregnated traps baited with a variety of chemical attractants, including fermented cow urine, to lure the tsetse. Insecticides applied to the backs of cattle also offer protection. Once the tsetse population is reduced to fairly low levels (often less than 1% of the original population) by trapping

and sometimes by large scale aerial spraying of insecticide, large numbers of male tsetse flies, factory reared and sterilized by x-ray strength gamma rays, are released. Since females mate generally only once in their life, mating with a sterile male will prevent reproduction. This strategy has been used to eradicate tsetse from the East African island of Zanzibar (Saleh *et al.,* 1997) and is useful for tsetse control in relatively small, defined areas where tsetse are not easily re-introduced.

1.4 Trypanosomes

Salivarian trypanosomes are the cause of both human African sleeping sickness and nagana. Salivarian trypanosomes, in contrast to stercorarian trypanosomes such as *Trypanosoma cruzi*, develop in the anterior part of the tsetse fly alimentary canal and are transmitted via the mouthparts. The name Trypanosome is derived from the Greek *trypano* (borer) and *soma* (body) because of their corkscrew-like motion. African trypanosomes are heteroxenous, salivarian protozoan parasites belonging to the Order Kinetoplastida, Genus Trypanosoma. Members of this genus alternate between two very different life cycle stages. Trypanosomes are unusual among the eukaryotes in that they have a specialized mitochondrion that exists in the form of a single elongated tubular structure (Figure 1.3; green) that extends along the length of the cell body. All the mitochondrial DNA is contained within the kinetoplast (Figure 1.3; red), a discrete structure located in a distended portion of the mitochondrion near the base of the flagellum, at the flagellar pocket (Robinson and Gull, 1991).

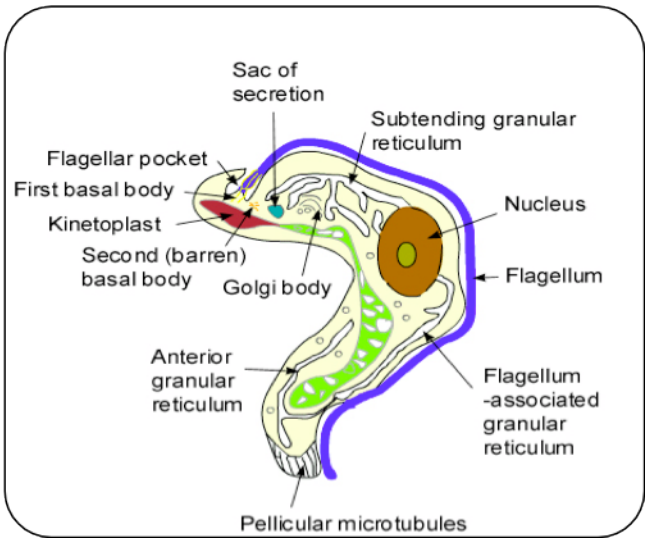

Figure 1.3. ***Schematic representation of the principal structures of the long-slender bloodstream form of the salivarian trypanosome, Trypanosoma congolense, revealed by electron microscopy.***
(Adapted from Vickerman, K., 1969. J Protozool. 16:54-69.)

A typical trypanosome is an elongate organism, 15 – 30 μm in length and 1.5 – 3 μm in width, with a single predominant nucleus (brown) containing a large central nucleolus, a complex endoplasmic reticulum and a single branched mitochondrion. Movement is effected by a single flagellum (blue) which is located at the posterior end and runs along the free edge of an 'undulating membrane' continuing anteriorly as a free flagellum. Locomotion occurs with the flagellum leading. Ultrastructure studies have revealed marked cytological and metabolic differences between the bloodstream forms and tsetse gut forms (Vickerman 1962, 1985). In the bloodstream there exists two major forms: long slender bloodstream forms and short stumpy bloodstream forms, with the former having a poorly formed "ghost" mitochondrion extending the length of the body, with few cristae present. The short stumpy bloodstream forms contain a developing mitochondrion with well-formed cristae. In contrast, in the tsetse gut forms, the mitochondrion takes the form

of an extensive tubular network with well-developed cristae. These structural differences are reflected in metabolic differences: slender bloodstream forms have a wasteful method of energy production that involves anaerobic breakdown of glucose to pyruvic acid, subsequently discarded, whereas short stumpy bloodstream forms as well as fly gut forms produce energy more efficiently through the aerobic breakdown of pyruvic acid (Vickerman, 1966).

1.5 Trypanosome life cycle

The life cycle of African trypanosomes is complex. These parasitic protozoa alternate between an insect vector, the tsetse (*Glossina* spp.) and a mammalian host. Both trypanosome stages are subject to dramatic changes in environment and therefore it is not surprising that their response, in terms of their metabolism (discussed previously) and surface architecture, is equally dramatic. Since my thesis is focused on surface molecules, I will discuss this aspect in some detail to set the stage for the research that is presented.

There are four major life cycle stages of African trypanosomes: bloodstream forms (BSF), procyclic forms (PF), epimastigote forms (EMF) and metacyclic forms (MF). Both BSF and MF are exposed to the host whereas the other forms are found only in the tsetse vector. Trypanosomes express different types of stage-specific surface proteins, with those so far described all anchored to the surface via glycosyl-phosphatidyl-inositol (GPI)-anchors. During all life cycle stages the trypanosomes are covered with a continuous monolayer consisting of proteins, glycoproteins or other glycoconjugates. BSF and MF are covered with a dense homogeneous coat of variant surface glycoprotein (VSG) that is involved in evasion of host immune responses. The life cycle stages that are found only in

tsetse express cell surface molecules that have been proposed to protect the parasite from proteolytic digestion (Acosta-Serrano *et al.,* 2001) or to serve in parasite development and possible ligand-associated parasite-vector signaling (Richardson *et al.,* 1988; Roditi and Pearson, 1990; Roditi *et al.,* 1998; Ruepp *et al*., 1997) or cell death (Pearson *et al.,* 2000). These two major groups of surface molecules will be discussed in turn below.

Mammalian blood infected with African trypanosomes contains extracellular BSF parasites that are covered with a dense, highly immunogenic surface calyx of approximately 10^7 identical VSG molecules (Vickerman, 1969). This continuous, dense monolayer of a single type of glycoprotein acts as a physical barrier shielding non-variant underlying membrane proteins from host immune responses. VSGs are central to antigenic variation, the phenomenon used by the trypanosome population to avoid elimination of the population from the infected host. In an infecting population, the consecutive and unpredictable expression of a series of VSGs from a large reservoir of approximately 1000 to 2000 VSG genes permits expansion of antigenically distinct trypanosome populations within the host. Different VSGs are antigenically distinct due to their extreme variation in sequence, but they also have an overall conserved structure, presumably indispensable for their function as a protective barrier (Blum *et al.,* 1993; Chattopadhyay *et al.,* 2005). The unfortunate consequence of antigenic variation is that there is little hope for a conventional anti-trypanosome vaccine being produced any time in the near future. Antigenic variation has thus produced both an immunological and intellectual barrier to vaccine development.

There are a few trypanocidal drugs in circulation, but the treatment of patients is problematic for numerous reasons. First, the drugs are toxic- the mortality rate from the drugs alone can be as high as 5 %, thus patients often require hospitalization for safe drug

administration. Second, the drugs are specific for either early or late stage disease, thus patients must be correctly diagnosed. Third, patients require regular check-ups (which is difficult in remote villages) to monitor relapse. Fourth, the efficacy of currently available trypanocidal drugs is thwarted by drug resistance that is developing faster than anticipated (Clausen *et al*., 1992).

BSF parasites rapidly proliferate as long, slender forms and when they reach a specific threshold density, short stumpy BSF begin to appear. This transformation appears to be a mechanism that pre-adapts the trypanosome for life in the tsetse vector. Not all long slender BSF differentiate into short stumpy forms, but those that do are eliminated from the host unless they are taken up by the tsetse. When tsetse consume infective bloodmeals, the VSG-covered BSF, along with the blood, enter the fly midgut. The non-proliferating short stumpy BSF (Figure 1.4; Panel A), which has a semi-developed mitochondrion with tubular cristae and activated proline oxidase and oxoglutarate oxidase systems (Vickerman *et al.,* 1988), seem to be pre-adapted for survival in this new environment that is cooler (27 ^{o}C) and proline-rich. In the midgut these pre-adapted BSF irreversibly differentiate into procyclic midgut forms (Vickerman *et al*., 1988) (Figure 1.4; Panel B). This transformation is characterized by loss of the VSG coat and expression of a new, more restricted set of GPI-anchored, tsetse-specific surface glycoproteins (Roditi *et al*., 1989). Procyclic forms have a functional mitochondrion and exhibit oxidative phosphorylation for generation of ATP using proline as the main energy source.

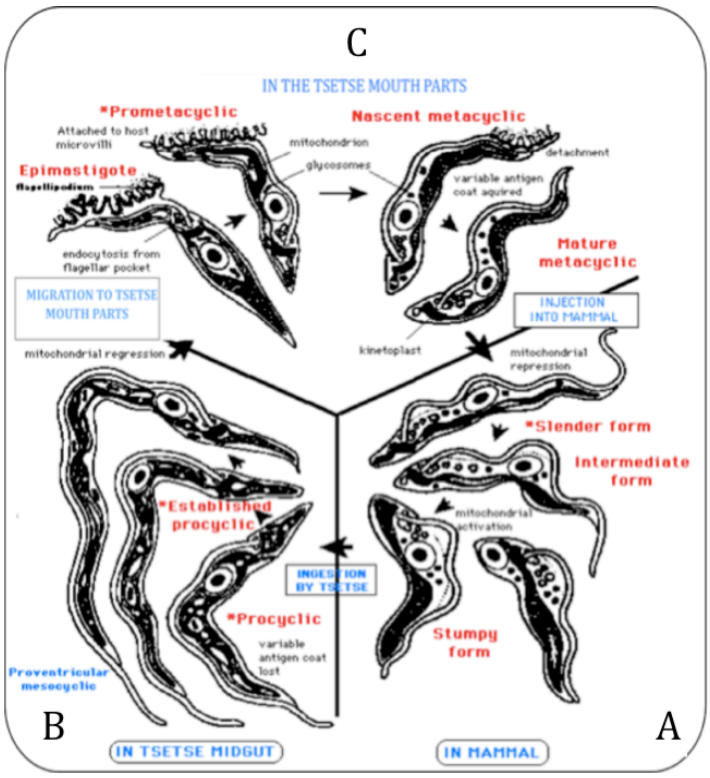

Figure 1.4. Schematic representation of the developmental cycle of African trypanosomes in host mammals and tsetse vectors.
In tsetse, the ingested short stumpy BSF transform to procyclic forms in the midgut. The procyclic forms migrate to either the salivary glands (Trypanozoon) or the proboscis (Nannomonas) where they differentiate into the adhering EMF followed by final transformation into the mammal-infective MF that express a VSG surface coat (modified from Vickerman, 1985).

Following successful survival and establishment within the tsetse midgut, the differentiated procyclic forms progress through two more major life cycle stages as they transit the tsetse from midgut to their final destination, either the proboscis or the salivary glands. Trypanosomes are defined in part by their developmental cycle in the tsetse. While EMF and MF trypanosomes of subgenus *Nannonomas* develop in the proboscis, the corresponding forms of subgenus *Trypanozoon* parasites develop in the salivary glands of tsetse. In the mouthparts, the parasites undergo their second morphological change, transforming into EMF (Figure 1.4; Panel C) that adhere to the surfaces of cells lining the

salivary glands or mouthparts via their flagella through hemidesmosome-like structures (Vickerman *et al.,* 1969; Evans *et al.,* 1979). A period of multiplication by binary fission follows. The EMF must further differentiate, in a process termed metacyclogenesis, into the infective MF which are mammal-infective, VSG-expressing, non-dividing and free swimming (Figure 1.4; Panel C) (Vickerman *et al.,* 1988; Hendry *et al.,* 1988*)*. This ends the cycle in the fly. Inhibition of attachment of the epimastigotes does not appear to prevent division of the parasites, but it does inhibit differentiation into MF, which implies that attachment of the epimastigotes has developmental significance (Vickerman *et al.,* 1988).

The period from ingestion of trypanosome-infected blood to the appearance of the mammal-infective metacyclic forms varies from one to three weeks. When an infected tsetse takes its next bloodmeal, these non-proliferating metacyclic parasites differentiate into proliferating BSF after being injected into a new vertebrate host, thus completing the trypanosome life cycle.

African trypanosomes have been studied extensively using a molecular approach, with research primarily focused on *T. brucei* spp. Metabolic pathways of the parasites have been studied to find drug targets and surface coat molecules of BSF and PF have been studied because of the perceived importance of host-parasite and vector-parasite interactions, respectively. More recently, the study of trypanosomiasis of livestock caused by *T. congolense* has picked up momentum, with particular attention being focused on the tsetse-infective forms, as these do not appear to demonstrate antigenic variation.

1.6 T. congolense tsetse-specific surface molecules

The best-known surface molecules of insect forms of trypanosomes are the procyclins expressed by *T. brucei* spp. (Roditi and Pearson, 1990). The procyclins are expressed on the procyclic surface at approximately three million GPI-anchored protein molecules per parasite. These unique molecules possess extensive glu-pro (EP) or gly-pro-glu-glu-thr (GPEET) repeats (Pays and Nolan, 1998; Roditi *et al.,* 1998; Roditi and Lehane, 2008) and exhibit distinct expression profiles *in vivo*. Within a few hours of differentiation into PF, both EP and GPEET molecules are expressed (Vassella *et al.,* 2001). However, three days post infection GPEET predominates with lower amounts of EP and from approximately the seventh day onwards, only the EP forms of procyclin are expressed (Acosta-Serrano *et al.,* 2001). The function of procyclins remains unknown.

T. congolense procyclic forms express several major stage-specific surface molecules that are different from those of *T. brucei*. In the early stage of a *T. congolense* midgut infection, a 24-34 kDa protease-resistant surface glycoconjugate (PRS) predominates (Bütikofer *et al.,* 2002). A second molecule, glutamic acid/alanine rich protein (GARP; Beecroft *et al.,* 1993) is weakly expressed or absent at this early procyclic stage (Bütikofer *et al.,* 2002). PRS and GARP are not easily detected in most established procyclic midgut forms. During this intermediate stage a 50-58 kDa heptapeptide (EPGENGT) repeat-containing molecule is expressed (Utz *et al.,* 2006). This *T. congolense* heptapeptide repeat protein (TcHRP) is thought to be the equivalent of the *T. brucei* procyclin molecules. Neither PRS nor the TcHRP are present on the surface of EMF in the tsetse mouthparts. In this stage, two different molecules predominate: GARP, (Bütikofer *et al.,* 2002) and a 100

kDa *congolense* epimastigote-specific protein (CESP) that is only expressed in epimastigotes (Sakurai *et al.,* 2008).

However, things are not always cut and dry and there is some protein 'wobble', whereby altered expression levels in different strains or cultures (e.g. *in vivo* versus *in vitro*) exist. *In vivo*, GARP is weakly expressed in early midgut PF, is down-regulated in established midgut parasites and strongly expressed in EMF isolated from the tsetse proboscis. *In vitro*, GARP is strongly expressed in procyclic culture forms (PCF) and epimastigote culture forms (Bütikofer *et al.,* 2002). This illustrates the importance of environmental factors on differentiation when using *in vitro* cultured forms, since they can exhibit altered regulation. Despite a lack of sequence similarity between tsetse-specific proteins of *T. congolense* and procyclins of *T. brucei*, common features such as surface-orientation, immunodominance and acidity are shared. Like the *T. brucei* procyclins, the surface molecules of *T. congolense insect forms* are lipid-anchored, show subgenus-specific distribution and are often glycosylated.

1.7 Protease Resistant Surface molecule (PRS)

Compared to the *T. brucei* procyclins, relatively little study has been conducted on *T. congolense* tsetse-specific surface molecules. PRS is an abundant *T. congolense* surface glycoconjugate that is expressed early in tsetse midgut infection and decreases during parasite establishment until it is completely absent in established midgut forms and in epimastigotes in the proboscis. Based on [^{3}H]ethanolamine labeling, high resistance to proteases and lack of amino acids, PRS was determined to be a GPI-anchored polysaccharide (Bütikofer *et al.,* 2002). Recent mass spectrometric analysis of the chemical

composition and GPI anchor structure of PRS showed that PRS has a unique lipid structure as well as an unusual content of carbohydrate residues (Greganova *et al.,* 2010). The lipid anchor of PRS, which is tri-acylated, has an unusual mixture of inositol-acylated diacyl-phosphatidylinositols, including species containing either myristic or oleic acid at the *sn*-2 position of the glycerol backbone. It is unknown whether there is a functional advantage to having a tri-acylated GPI anchor. Analysis of the monosaccharide content of PRS by gas chromatography-mass spectrometry (GC-MS) further revealed PRS to be a GPI-anchored glycoconjugate with an approximate compositional ratio of Man: Glc: Gal: sialic acid (SA) of 1: 1.7: 21.6: 3.3 (Greganova *et al.,* 2010). Although trypanosomes are unable to synthesize sialic acids, they make use of a trans-sialidase to transfer sialic acid from the environment onto trypanosomal surface molecules (Tiralongo *et al.,* 2003). PRS is highly rich in galactose residues, and has a relatively high content of sialic acid (SA), suggesting that PRS may be a major acceptor of sialic acids in *T. congolense* procyclic forms. In *T. brucei*, sialylation of the procyclin anchor has proven to be essential for parasite survival within the tsetse midgut (Nagamune *et al.,* 2004). In *T. brucei* spp. and *T. congolense*, transialidase is only expressed in the procyclic insect stage. Using this enzyme at this stage to sialylate the major cell surface glycoproteins (e.g. EP procyclins in *T. brucei* spp., and PRS and GARP in *T. congolense*) confers a negatively charged glycocalyx to the parasites, which is believed to protect them from digestive conditions in the fly gut, or from immunoactive substances present in the tsetse blood meal (Tiralong *et al.,* 2003).

1.8 T. congolense Heptapepeptide Repeat Protein (TcHRP)

In 2002, researchers in the labs of Peter Bütikofer and Isabel Roditi in Bern, Switzerland, set out to identify novel *T. congolense* surface components by labeling surface molecules using radioactive GPI precursors. These groups had noticed that unexpectedly, 14 days post-infection, procyclic forms were frequently negative for both PRS and GARP, suggesting that another stage-specific surface molecule was expressed at this point in the life cycle. Indeed, in addition to PRS and GARP, *T. congolense* procyclic culture forms were found to express a third, 50-58 kDa trypsin-resistant GPI-anchored molecule (Bütikofer *et al.,* 2002). To investigate whether the latter might constitute the coat of these parasites the protein was purified and used to generate a polyclonal antiserum and the protein's encoding DNA was sequenced. The antiserum was used in immunofluorescence experiments to examine trypanosomes taken directly from infected tsetse, showing that the molecule was surface expressed (Utz *et al.,* 2006). A DNA sequence encoding a protein with 13 identical EPGENGT heptapeptide repeats was found. This *T. congolense* heptapeptide repeat protein (TcHRP) showed extensive N-glycosylation, with at least 10 (but up to 13) high-Man-type oligosaccharides ranging from $Man_5GlcNAc_2$ to $Man_9GlcNAc_2$ and contained a GPI-anchor with potential complex GPI modifications (Utz *et al.,* 2006). Additionally, the threonine (Thr) residues in the heptapeptide repeats were occasionally modified with oligosaccharide chains linked via phosphodiester bonds. Taken together, TcHRP has a much higher apparent molecular mass (50-58 kDa) than predicted solely from the amino acid sequence (8.1-9.5 kDa). Similar observations have been made for GARP (Bayne *et al.,* 1993; Beecroft *et al.,* 1993; Bütikofer *et al.,* 2002) and the *T. brucei* EP and GPEET (Clayton *et al.,* 1989; Field *et al.,* 1991; Bütikofer *et al.,* 1997) procyclins. It is thought that the

aberrant apparent molecular weights of these molecules based on gel electrophoresis are due to their complex carbohydrate-glycolipid anchors, to glycosylation of the main polypeptide backbones and to the high acidity of the polypeptides that repels SDS thus altering properties of charge:mass ratio and, subsequently electrophoretic migration. The GPI lipid portion of TcHRP has been reported to consist exclusively of inositol-acylated *lyso*-acyl phosphatidylinositol species (Utz *et al.,* 2006). Much like the *T. brucei* procyclins, TcHRP is expressed continuously during the course of a midgut infection. It is of interest that the N-terminal domain of the molecule isolated from parasites grown in tsetse is shorter than predicted from its encoding DNA sequence. During infection in the tsetse midgut, the N-terminal domains of EP and GPEET are quantitatively removed by proteolysis, leaving a sequence of protease-resistant amino acid repeats (Acosta-Serrano *et al.,* 2001). These repeat domains of procyclins of both *T. brucei* spp and *T. congolense* are thought to form a shield that protects susceptible proteins on the cell surface from tsetse gut proteases that are especially active during digestion of a blood meal. Because of these shared characteristics with *T. brucei* EP and GPEET procyclins, TcHRP is referred to as the *T. congolense* procyclin.

1.9 Glutamic acid/Alanine-Rich Protein (GARP)

In 1993 two labs independently discovered the first non-VSG surface molecule expressed on *T. congolense* (Beecroft *et al.,* 1993; Bayne *et al.,* 1993). Monoclonal antibodies (mAbs) were raised against lysates of *T. congolense* PCF and those that bound to the surface of living parasites in immunofluorescence experiments were selected. The selected mAbs bound to living PCF of *T. congolense* and *T. simiae* but not to *T. b. brucei, T. b.*

rhodesiense, or to *T. vivax*. They also bound to *T. congolense* EMF but not to MCF or BSF (Beecroft *et al.,* 1993). These antibodies were thus life cycle stage- and subgenus-specific. All antibodies detected broad bands of 40-44 kDa (GARP) and 28-32 kDa (PRS) in immunoblot analysis of procyclic lysates and they were specific to carbohydrate epitopes. The 40-44 kDa antigen was purified and sequenced (Beecroft *et al.,* 1993; Bayne *et al.,* 1993). Amino acid microanalysis of the purified protein revealed a high content of glutamic acid, glutamine and alanine. A gene encoding a protein of 256 amino acids with a potential N-terminal leader sequence and a C-terminal hydrophobic sequence was identified. The protein was termed glutamic acid/alanine rich protein (GARP) (Beecroft *et al.,* 1993; Bayne *et al.,* 1993). GARP is highly glycosylated, containing large galactose-rich phosphoglycans linked to the protein backbone and like TcHRP, GARP is GPI anchored with a GPI lipid portion consisting exclusively of inositol-acylated *lyso*-acyl phosphatidylinositol species (Thomson *et al.,* 2002). GARP has a much higher apparent molecular mass (40 – 44 kDa) than predicted (21.6 kDa) based on the amino acid sequence. *GARP* genes are present in all trypanosomes of subgenus *Nannomonas*, however, based on nucleotide and amino acid sequence comparisons, GARP has diverged considerably within the subgenus (Asbeck *et al.,* 2004).

1.10 Congolense Epimastigote-Specific Protein (CESP)

EMF in tsetse mouthparts and salivary glands have an adhesive phenotype. The molecular basis of epimastigote adhesion has not been well studied due to both the difficulty in obtaining sufficient numbers of parasites from either the proboscis or salivary glands, and the absence of a model for studying adhesion (Vickerman, 1969; Evans *et al.,*

1979; Thevenaz *et al.,* 1980; Beattie *et al.,* 1997). Investigators had concluded that attachment of the epimastigote was a necessary and essential part of the developmental program for metacyclogenesis (Hendry *et al.*, 1988). In 2008, Sakurai and coworkers identified a novel stage-specific GPI-anchored surface glycoprotein that was selectively expressed in the epimastigote life cycle stage of *T. congolense*. They named it *congolense* epimastigote-specific protein (CESP) (Sakurai *et al.,* 2008). Using an *in vitro* culture system, they were able to obtain a sufficient amount of epimastigote material to allow identification and characterization of CESP. CESP was predicted to be a highly α helical protein with an internal region consisting mainly of hydrophilic amino acids, but with termini having approximately 20 hydrophobic amino acids. The hydrophobic N- and C-terminal domains were predicted to be a signal peptide and GPI-anchor binding region, respectively. The *CESP* gene was present in multiple copies and encodes a polypeptide of 689 amino acids. The amino acid composition indicated that CESP has six potential asparagine residues for N-glycosylation (N-XS/T). Much like GARP and TcHRP, CESP has an apparent molecular mass higher (100 kDa) than that predicted (72.9 kDa) from the translated amino acid sequence, suggesting that CESP is highly glycosylated.

Molecular characterization of the four surface molecules in *T. congolense* insect forms has been mostly achieved, although their complete structures have not been determined and many questions remain about their temporal expression throughout the *T. congolense* life cycle. Their functional roles remain unknown.

1.11 Purpose of this thesis

Although culture systems have improved dramatically over the years and trypanosome isolates have been successfully cultured in the lab, only with *T. congolense* can all four main life cycle stages (BSF, PCF, EMF and MCF) be grown *in vitro.* Trypanosomes grown *in vitro* and *in vivo* appear to be essentially identical as determined by similar ultrastructural characteristics and their utilization of identical respiratory substrates (Vickerman *et al.,* 1988). In addition, high-resolution, two-dimensional polyacrylamide gel electrophoresis analysis revealed that the proteins expressed by the PCF and the PF taken directly from the tsetse midgut were qualitatively similar (Pearson *et al.,* 1987). Thus *T. congolense* is amenable to biochemical studies, particularly since all life cycle stages can be grown *in vitro* in sufficient amounts allowing protein chemical and immunochemical analysis.

The aim of my thesis was to generate specific monoclonal antibodies to three of the four *T. congolense* tsetse-specific surface molecules: GARP, TcHRP and CESP, each of which has a polypeptide backbone. I believe that such antibody probes will be useful for studying the expression of these markers throughout the life cycle of *T. congolense* and for studying the immunochemical structure of the molecules themselves. The information thus gained will facilitate the understanding of the structure-function relationships of these major surface molecules in the tsetse vector.

Chapter 2. General Materials and Methods

2.1. Trypanosomes

All four life cycle stages of *T. congolense* IL 3000 (Fish *et al.,* 1989) were used for most of the work presented in this thesis. In addition, PCF of *T. congolense* K45/1 (derived from *T. congolense* STIB744; Swiss Tropical Institute, Basel, Switzerland), *T. simiae* CP-11 (Zweygarth *et al.,* 1987; Majiwa *et al.,* 1987) and *T. b. brucei* 427-01 (Cross and Manning, 1973) were used.

The four major life cycle stages of *T. congolense* clone IL3000 were prepared as previously described (Helm *et al.,* 2009). Since these life cycle stages are central to the work reported in this thesis, I will describe their preparation in some detail. Together with my collaborators, Dr. Noboru Inoue and Dr. Tatsuya Sakurai at Obihiro University, Obihiro, Hokkaido, Japan, we prepared some of the trypanosome life cycle stages that I used. My collaborators also taught me to grow several of the insect stages *in vitro*.

BSF stabilates of a cloned population of *T. congolense* IL3000 (Savannah group) (Bienen *et al.,* 1991; Fish *et al.,* 1989) were stored in liquid nitrogen at the National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido, Japan. To grow BSF, frozen stabilates were thawed and aliquots inoculated intraperitoneally into five, 8-week-old female C57BL/6 mice. At the first peak of parasitemia, the mice were bled by cardiac puncture and trypanosomes were purified from the infected blood using DE-52 cellulose (Whatman, Middlesex, UK) anion exchange column chromatography (Lanham *et al.,* 1970). All animal experiments were

performed according to the standards for Care and Management of Experimental Animals at Obihiro University of Agriculture and Veterinary Medicine. No animals were used for growing trypanosomes at the University of Victoria. Cultures of PCF, EMF and MCF were produced *in vitro* from the IL3000 BSF grown *in vivo* by following the methods of Hirumi and Hirumi (1991). BSF cultures were adjusted to 3 x 10^6 cells/ml in Eagle's minimum essential medium containing 20 % heat-inactivated fetal bovine serum, 2 mM L-glutamine and 10 mM L-proline. This suspension (10 mL) was incubated in 25 mL culture flasks at 27 °C for approximately one week to allow differentiation to PCF. The PCF were maintained in log-phase growth for several weeks to allow complete differentiation. For EMF production, cultures of PCF were allowed to become slightly acid. After one to two months, adherent clusters of parasites began to appear on the flask surfaces. When the adherent cells formed a monolayer, they were washed with sterile phosphate-buffered saline 1 % glucose, pH 7 (PSG) to remove non-adherent parasites before harvesting EMF. These parasites were identified by light microscopy to be EMF. Only a few contaminating PCF were seen. After a few weeks of incubation of flasks with confluent parasites, non-adherent, VSG-expressing MCF began to appear in the supernatant. To obtain the various life cycle stages used in my work, different cultures were used. PCF were collected from suspension cultures of PCF growing in log-phase. EMF were collected from flasks containing confluent monolayers by first gently washing away loosely bound cells (dead cells, PCF and MCF) using three washes of 10 mL PSG followed by gentle removal of adherent cells from the culture flasks. These were resuspended in PSG. MCF were purified from EMF culture supernatants by anion exchange chromatography as described above for BSF. Finally, the purified MCF were used to inoculate BALB/c mice (10^5 parasites per animal)

intraperitoneally. Approximately seven days after infection, the infected mouse blood was collected by cardiac puncture and the BSF purified using DE-52 column chromatography. *T. congolense* IL 3000 PCF were routinely maintained by diluting 5 mL of log-phase parasite suspension with 10 mL of fresh medium every day. The plastic-adherent EMF cultures were maintained by replacing the entire culture supernatant with fresh medium every five days. All PCF and EMF were maintained in culture at 27 ^{o}C in a modified minimum essential medium (MEM) containing Earle's salts/ 25 mM HEPES/ 10 % heat-inactivated fetal bovine serum (FBS)/ 1 % non-essential amino acids/ 2 mM glutamine/ 2.6 μg mL^{-1} hemin/ 60 μM proline and 200 μm hypoxanthine (PCF medium; Fish *et al*., 1989). In addition, 50 units mL^{-1} of penicillin-streptomycin were added to the medium.

2.2. Recombinant proteins

Recombinant GARP: Dr. Noboru Inoue generously provided recombinant GARP-GST fusion protein that was used for some of my work. A cDNA expression library was produced from *T. congolense* IL 3000 Savannah type, isolated near the Kenya/Tanzania border in 1966 (obtained from the Biological Services Unit at ILRI, Nairobi, Kenya) and the GARP sequence was PCR-amplified before expression. Briefly, a partial *GARP* sequence (100-666 bp) was PCR-amplified using a forward primer containing a BamHI site (underlined) (5' –GGATCC **CAG AGC GTT CCC CCA AAG GT**-3') and a reverse primer containing an *EcoRI* site (underlined) (5' – GAATTC **GGC CTT CTC CGC CTC GTA CT**-3'). The PCR products were subcloned into a pGEX4T-1 plasmid vector in *Escherichia coli* (*E. coli*) DH5α and expressed as a glutathione S-transferase (GST)-fusion protein. Recombinant

protein was purified using glutathione Sepharose 4B beads and sent to us at a final concentration of 2 mg GARP-GST/mL (Sakurai *et al.,* 2008).

ATGACGACAACCATGTCCCGTGTCCTACACCTAATGACGGTCACACTGCTGTGCGCACGCGTGGG
AATGGGCCAGGCTAGCGATGACGACGATTGTGGCGGG**CAGAGCATTCCCCAAAAGGT**GGAGGAG
GTGCAAACGATGTGCGACGTTGCGCGGCAGCTGAGGGCCCTGGAGACCGCTTCCCAGTCCGCGGT
GGGTGCCGTGGTTTCTTCTGCCCGGGATGCGTCGGAGGCGAAAGAGCGTGCGGAGAAAGCTGTGG
AGCGCGCCAAATCGAAGAAGCGTGGTGTGGACGCGGCGACGGAAGCGGCTGCAAGGGCTGCGGCT
GCGGCCCAGCGCGCGGAGACGGTGGTGAGCGATGCGAGGAAGCACGCGGCAGACCTGACGGCGGC
ATCGAAGGATGCTATCGAGACGACCGACGAGTCGCTGCGCCTACTGGCCACATGCGAGAAAGCGG
ACGAGCCCATCCGCACTGCTGCAAAAAAGTGCACGGGTGCCGCCGCCGAAGTCACGTCCAAGTCCC
TTGAGTCAGCGTTCGACGCTCTCGCGGAACTGCTACCGGATGGTGCGGACGACATCCGCGAGCAC
GGTGCCGTGTTCGTGAAGGGGCTGAAGTCTCTGGAGGATGACGTGCGCACGGCTGGAGAGGCAAA
GAGC**GAGGCGGAGAAGGCT**GAGGGCGATGCGAACGACGCGGCAGATGGTGCCCGTGCCGTGCTG
ACGGGCGTGTGCGTGCTGCTGCTTCTGGCTGCACTGCACTTTTCTGCGGGGCTGTGA

Figure 2.1. GARP nucleotide sequence: TcIL3000.0.60820.
Shown between the PCR-amplified forward and reverse primers (bold) is the partial GARP DNA sequence (100-666bp) that was cloned into the expression vector.

Recombinant CESP: Dr. Inoue also provided some of the recombinant CESP-GST fusion protein used in my work. His lab constructed and screened a cDNA expression library from *T. congolense* IL 3000 Savannah type, isolated near the Kenya/Tanzania border in 1966 (Biological Services Unit at ILRI, Nairobi, Kenya). A partial *CESP* sequence (61-2007 bp) encoding a hydrophilic region of the protein was PCR-amplified using a forward primer containing a BamHI site (underlined) (5' –GGATCC **AGT GGC GAT GAG TTC TCC TT**-3') and a reverse primer containing an *EcoRI* site (underlined) (5' – GAATTC **ACC GTG TCG CGC ACC GCC** AG-3'). The PCR products were subcloned into a pGEX4T-1 plasmid vector in *Escherichia coli* (*E. coli*) DH5α and expressed as glutathione S-transferase (GST)-

fusion protein. Recombinant protein was purified using glutathione Sepharose 4B beads and sent to us at a final concentration of 2 mg CESP-GST/mL (Sakurai *et al.,* 2008).

GGATCC**AGTGGCGATGAGTTCTCGTT**GGATCATGTAAACGCATTCTGTGAGCTGACGAAGCAGT
TTAGGCATTTGCCAATCGCCGTGGGGCAGCAACTTGATGAGGCAATAGATGAGGAGGCCAAGGCC
GCGAAAGCGAGGAAAGATAGCCAAAAGGCTGTGAACCGTGCCGATGCCGCCGCGAAGAAGAGCGA
GGATGCCAAGAAGCACGCCCAAGATGCAAAGGAGGCAGCCAAGGAGGCATTTGAAGCAGCGGACG
GGGCTGAGGTTCATTTACAGGCGGCTTCTGAACTAGCGACTGAGCTTAACGGACTGGTTGAAACG
CATCTATCAAGGCTTGAAGAGTATCTCAAGGATGCAGAAACCAAAGGAGAGGACGAGCGTGCCCA
GGAAGCCGCGAAGGAGTGCACTGGGAACGCCACGAACGTTACCTCAGGCACGTTGCTGGAATCCA
AAGAGAAACTGGAGAAAGCTGCGGGTAAGGAAGACTACGAGGAATTGAGGTATGACACGGAGCG
GGCTGGCTCGCTGCTGGACGACCTCAAGGGTGCCCAGCTTGAAGTGTCTTCGCTCCAGATTAAGG
CTACCAATGCCGAGGAGCGTGCGACGGAGGCTGCTGCGAAGGCGAAGGCCGCCACTCCCGATTTA
TACGATCTCACGGTAGGACAGGCGAAAGCGTTTTGCAAACTAACTGAGCAACTGCGTGGACTGCT
GGATACTGTATCGAGGCACAATGTCACGGTGACGGAGGAAGTGGCGAAGACTGCATCCTCGAAGA
ATCGTAGTGATGAGGCCGTGAAGCAGGCTGAATCCGCTGCCGGTAGGAACAGTGATGCTGCACCC
CATGCGAAGAAGGCGAAGGAGGCTGGTGAGGAGGTTTCTGCCGCTGCTGACGAGGCGCAGAAAGC
ACACGAGGACAGTAGCAGAGTGGTGCAGGACCTGAGATTGCTGGCCAAAGACAAGCTTTCCTCCT
ATGACCGGTTTCTGAAGAATGTCGGTGAGCGTGCAAGCGGAAAAGATGCTCAAAACGCTGCGAAG
GAGTGCACGGATAGCGCAGTGGATGTGAAGTCCGAGGTGGTGGAGAAGCTCAATGCAGCGTTCGA
GGAGAGCACCCCTGGCAAGCCGTTTCAGCCTGTGAAGGATCTGTTCAAGAAATTTTCTGACAACC
TCAAGGAGCTGGAGGGTGCTGCGGATGTGGCATCCAAGTCTCGTGCCAAGGCCGAGGCTGCCGAG
AAGATCGTGAACGAGTCCGCTGCTGAGGCTGAGATTGCTGCCTTTGATGTGAAGAAGCTCACGTT
GAAGCAGGTGAACGCATTCTGTGGGCTGGCGGAGCAGTTCCGCGGGCTTTTGAACTCCGTGAAGA
AGCTGGAAGGCCAGGCGACGCACTGTGCAACGGTGGCTGCGGCAGTAAAGGCGCGAAGCACGGAA
ACAGTGAAGCGCGTGGTGGCTGCAGCCGGGAAGAATGCGAGCGCGAAGGAACTTGCTGAGCGTGC
ACAGGCTGTTGAAGCTGAGGTGACTGCGGCGGCTGACAAAGCACAGGAGGCGTATGCGGGTGTCT
CTGGTTTCGTGGGGAATCTGACGGCGTCGAGGGACGAGCACCTCCTCGTGCTTGAGGATGTGATC
ACAGGGGCGAAGACGTCTGCATCGTGTAGCTGTGGTATTCGTGCTGCGAAGGCATGCACCGCTGC
CGCAGAGGACGTGACGGCGGAGTCTCTGAAGGAGGCCAAGGATACGTTCAGGCAAGACATCCCCG
GGGACCACTACAAGGAGCTGTCGGACGGCGCCGAAAATGTGTCTGGGCGTCTGGAAGAGCTGAGA
AGCACGGCGCAGCGCGTGTCTGCTTCACGGGCGGCAGCGCAGGCTGCTGAGGAGCGATTGAACGC
CATAGCCGAGGAGGCCGGTGTGAAACCTGCGGAAGAGGAGGGTGGTGCGG**CTGGCGGTGCGCGA**
CACGGTGAATTC

Figure 2.2. CESP nucleotide sequence: accession no. AB300788.
Shown between the PCR-amplified forward and reverse primers (bold) is the partial CESP DNA sequence (61-2007 bp) that was cloned into the expression vector.

2.3. Derivation of monoclonal antibodies

Six female BALB/c mice, aged six to eight weeks, were used for immunizations. Two mice were immunized with recombinant GARP-GST, two with recombinant CESP-GST, and two with *T. congolense* IL 3000 PCF lysates. All immunizations were performed over a four-month period. The first injection was administered intraperitoneally (IP) in Freund's complete adjuvant, composed as follows: 25 µL (2 mg/mL) recombinant protein/ 275 µL PBS/ 300 µL Freund's complete adjuvant, or 5×10^7 trypanosomes/ 200 µL PBS/ 200 µL Freund's complete adjuvant. Two subsequent IP injections were given in Freund's incomplete adjuvant composed as follows: 25 µL (2 mg/mL) recombinant protein/ 275 µL PBS/ 300 µL Freund's incomplete adjuvant, or 5×10^7 trypanosomes/ 200 µL PBS/ 200 µL Freund's incomplete adjuvant. The final boost of 25 µL (2 mg/mL) recombinant protein/ 375 µL PBS, or 5×10^7 trypanosomes/ 400 µL PBS was given intravenously (IV) three days prior to sacrifice of the mice. Each inoculum contained 25 µg recombinant protein per mouse, or 2.25×10^7 trypanosomes per mouse.

Three days after the last injection, mice were sacrificed by asphyxiation followed by cervical dislocation. The heart blood was collected (1 mL) and spleens were removed. The blood from the cardiac bleed was allowed to clot at room temperature for 1 hour, placed in the fridge at 4 ^{o}C overnight and centrifuged at 7,000 rpm (3615 x g) for 20 min on a SIGMA 1-14 microcentrifuge (Osterode, Germany). Serum, was saved to use as a positive control. The spleens were aseptically removed to a Petri dish containing 10 mL serum-free D-MEM medium (D-MEM; cat #SH30003.03/ 25 mM HEPES/ 44 mM $NaHCO_3$/ 2.86 mM 2-mercaptoethanol/ 2 mM L-glutamine/ 50 units mL^{-1} penicillin-streptomycin/ 2 mM sodium pyruvate), and teased apart to release splenocytes. The splenocyte suspension was

strained through a BD Falcon™ cell strainer (#352340, 40 μm nylon, BD Bioscience, Bedford, MA, USA) into a 50 mL conical screw-top tube, topped up to 50 mL with serum-free medium and centrifuged at 200 × g for 6 min on a Beckman TJ-6R, TH-4 rotor with buckets (cat no. 339235, Palo Alto, CA, USA). Supernatant was removed and the pellet was resuspended in 3 mL serum-free medium at room temperature. The suspension was slowly layered onto the top of 4 mL of room temperature Ficoll-Paque™ plus (#17-1440-02, Amersham Biosciences, Uppsala, Sweden) and centrifuged at 200 × g for 6 min, then 12 min at 650 × g. Using a sterile glass Pasteur pipette the hazy interface containing isolated lymphoctyes was removed from the Ficoll-Paque plus gradient and transferred to a sterile 50 mL conical screw-top tube. The purified spleen lymphocytes were topped up to 10 mL with warmed serum-free medium, resuspended and 50 μL removed to make a 1/10 dilution for counting on a haemocytometer. The SP2/O parental myeloma cells were counted to calculate the number of SP2/O cells needed to give a 1:5 ratio of SP2/O: lymphocytes. The appropriate volume of SP2/O and lymphocytes required was removed and separately transferred to a 50 mL conical screw-top tube, then topped up to 50 mL with serum-free medium and centrifuged for 6 min at 200 × g. Supernatant from both was decanted, pellets were loosened and the wash repeated. Loosened pellets and SP2/O parental myeloma cells were transferred to the tube containing the spleen lymphocytes. The cells were topped up to 50 mL with serum-free medium, mixed gently, and centrifuged for 5 min at 200 × g. Supernatant was decanted and all traces of medium removed. The pellet was loosened and the tube placed in a portable 37 ^{o}C water bath. The pellet was stirred continuously while 1mL of polyethylene glycol 1500 (PEG 1500) (Cat No. #783 641, Roche Diagnostics, Mannheim, Germany) was added over 1 min, followed by 1 mL, 3 mL, and 10 mL serum-free

medium over 1 min, 3 min, and 2 min, respectively. The reaction mixture was incubated for 5 min at 37 ^{o}C, then removed from the water bath, topped up to 40 mL with serum-free medium, to which 10 mL heat-inactivated fetal bovine serum (FBS) (#SH30396.03, HyClone Laboratories Inc., Logan, Utah, USA) were added. The fused cells were centrifuged for 6 min at 200 × g, supernatant removed and the pellet very gently resuspended. Cells were diluted in 5 mL Recovery medium (30 mL serum-free D-MEM/ 10 mL FBS/ 10 mL SP2/O fresh conditioned medium/ 0.5 mL IL-6 (100✕)/ 20 µL gentamicin/ 50 µL OPI stock (100✕). The fused cells were added drop-wise to a 25 cm^2 non-TC treated flask containing 50 mL recovery medium and incubated at 37 ^{o}C overnight. Single step selection and cloning of hybridomas was performed using the ClonaCell-HY system (StemCell Technologies Inc., Vancouver, B.C.). The cell fusion mixture was transferred to two 50 mL conical screw top tubes and centrifuged for 6 min at 1.67 × g. Supernatant was decanted from both tubes back into the 25 cm^2 flask, cells pooled and diluted in MethoCult semi-solid methylcellulose medium (MethoCult® M03134, STEMCELL technologies) containing 1 mL (100✕) HAT selective medium and B-cell growth factors: 1 mL OPI (100✕)/ 1 mL IL-6 (100✕)/ 1 mL Glutamax-1/ 20 µL Gentamicin/ 20 mL FBS/ 36 mL serum-free medium. Cells were incubated in a 37 ^{o}C, 5 % CO_2 incubator for at least 1 hour prior to plating. For plating, 9 mL of the hybridoma-containing MethoCult medium were aliquoted to each of twelve 60 x 15 mm Falcon® tissue culture Petri dishes (#353002, Becton Dickinson Labware, Franklin Lakes, NJ, USA). It was important to avoid introducing air bubbles. Six more 60 x 15 mm Falcon® tissue culture Petri dishes, with their lids removed, were filled with dH_2O. Two MethoCult-filled and one dH_2O-filled Petri dishes were placed into each of six 150 x 15 mm Petri dishes. These were then placed, lid on, into a 5 % CO_2, 37 ^{o}C incubator for growth of

clones. Approximately ten days later, 1000 clones were picked and resuspended in wells of 96 well tissue culture microplates in 200 µL of clone picking medium (Dulbecco's Modified Eagles/high glucose powder (D-MEM; cat #SH30003.03)/ 25 mM HEPES/ 44 mM $NaHCO_3$/ 2.86 mM 2-mercaptoethanol/ 2 mM L-glutamine/ 50 units mL^{-1} penicillin-streptomycin/ 2 mM sodium pyruvate/ 2 mM GlutaMAX-1/ 1X HT supplement (100X)/ 0.30 µg mL^{-1} IL-6/ 10 µg mL^{-1} gentamicin/ 20 % foetal bovine serum).

2.4. ELISA-based screening, selection and isotyping of monoclonal antibodies

Hybridoma tissue culture supernatants were screened by indirect ELISA on the immunizing antigen. ELISA plates (96 well Microlon 600™ Greiner flat bottom, high bind, chimney style, cat # 655081) were coated by drying 0.1 ml per well of: rGARP (1 µg/ 100 µL dH_2O), rCESP (1 µg/ 100 µL dH_2O), or *T. Congolense* IL 3000 PCF lysates (1×10^6 cells/ 100 µL dH_2O) overnight at 37 ^oC dry incubator. Screening of hybridoma supernatants was also performed on human transferrin (0.25 µg human transferrin (ht)/ 100 µL dH_2O) to detect non-specific "sticky" antibodies. Following antigen coating, 200 µl 3 % (w/v) powdered milk (#232100, Difco™ Skim Milk, BD and company, Sparks, MD, USA) in PBS was added to each well and plates were incubated for 1 hour at 37 °C. Plates were first rinsed (three quick washes using PBS-0.05 % Tween-20 pH 7.4), then primary antibody was added (100 µL neat tissue culture supernatants of the test mAbs) to each well, and incubated for 1 hour at 37 °C. After the first antibody incubation, plates were washed three times for 15min each with PBS-0.05 % Tween. Subsequently, 100 µL of secondary antibody was added (alkaline-phosphatase-conjugated goat anti-murine IgG/IgM; Cat No. #31328, Pierce Chemical Company, Rockford, IL; at 1/15,000 dilution in PBS-0.05% Tween/1 % skim milk)

and plates were incubated for 1 hour at 37 °C. Plates were washed as before, and then 100 µL of the substrate solution (one p-nitrophenyl phosphate pill; Sigma Chemical Company, Mississauga, ON, Canada; per 5 mL diethanolamine) were added to each well and the absorbance at 405 nm was read at 2 hours and again after overnight incubation at room temperature using an automated EIA plate reader (Model EK 310, Bio-Tek Instruments, Inc., Burlington, VT, USA). Positive hybridoma clones were expanded into 24 well tissue culture plates in clone picking medium (2.0 mL per well).

After stable hybridoma clones were established, isotyping of the mAbs was performed with 100 µL of tissue culture supernatants using an antigen capture ELISA technique. A polyspecific goat anti-mouse immunoglobulin antiserum was used as the capture antibody (Caltag, South San Francisco, CA, USA) and 1:5,000 dilutions of peroxidase-conjugated affinipure goat anti-mouse IgG1/ IgG2a/ IgG2b/ IgG3/ IgM (Jackson ImmunoResearch Laboratories, Inc., Baltimore, MD, USA) antisera were used as the isotyping reagents. The substrate used was 1-Step™Turbo TMB-ELISA (#34022, Pierce Chemical Company, Rockford, IL). Absorbance at 405 nm was read using an automated EIA plate reader (Model EL 310, Bio-Tek Instruments, Inc., Burlington, VT, USA).

2.5. 1-D SDS-PAGE and immunoblot analysis of monoclonal antibodies

One-Dimensional sodium dodecyl sulphate-polyacrylamide gel electrophoresis (1-D SDS-PAGE) was performed according to Laemmli (1970) using a Mini-Protean II minigel apparatus (BioRad Laboratories, Hercules, CA, USA). The following samples were prepared and resuspended to 20 µL in 1x Laemmli sample buffer, and boiled at 90 °C for 10 min on a Standard Heatblock (#13259-030, VWR Scientific products, USA): rGARP protein (0.5 or 1

µg/well), rCESP (0.5 or 1 µg/well), *T. congolense* IL 3000 lysates (2 × 10^6 cells/well), *T. congolense* K45/1 lysates (2 × 10^6 cells/well), *T. simiae* CP-11 lysates (2 × 10^6 cells/well) and *T. b. brucei* 427 lysates (2 × 10^6 cells/well). Twenty microlitres of sample were loaded into each well of a 0.75 mm, 3 % stacking gel with either a 10 % or 12.5 % resolving gel. Proteins were electrophoresed at 50 V for 30 min and then at 90 V for an additional 90 min. A PageRuler™ pre-stained protein ladder (10 – 170 kDa, #SM0671, Fermentas Life Sciences, Harrington, ON) was run on each gel. Following electrophoresis, gels were transferred onto polyvinylidene difluoride (PVDF) membranes (0.45 µm pore size, Immobilon-PTM, Millipore, Corp., Bedford, MA, USA) for subsequent immunoblotting for 30 min at 90 volts. Antigens were subsequently detected by the relevant test mAbs (1/1 dilution in blocking buffer: 5 % w/v skim milk powder in PBS-0.1 % Tween 20) or test bleed (1/200 dilution), followed by 1: 10,000 dilution (in blocking buffer) of horseradish peroxidase-labelled anti-murine IgG/IgM (H+L chain) specific secondary antibody (M30807, Caltag Laboratories, Burlington, CA). Control primary antibodies used were the following: 1. mAb TC 491 (IgG_3), which is specific for carbohydrate epitopes present on several molecules on the surface of *T. congolense* (Beecroft *et al.,* 1993; Mookherjee and Pearson, 2002; Loveless and Pearson, unpublished), 2. mAb 1-G1 (IgG_1), specific for subgenus *Nannomonas* (Mookherjee and Pearson, unpublished), and mAb 247 (IgG_1) specific for EP-repeat procyclin on *T. brucei* spp. (Richardson *et al.,* 1986; Richardson *et al.,* 1988). Briefly, once SDS-PAGE-separated parasite proteins were transferred, the PVDF membranes were incubated at room temperature for 1 hour in blocking buffer. PVDF membranes were washed three times in washing buffer (PBS-0.1 % Tween 20) for ten min each and then incubated overnight at 4 °C with the primary test antibodies. The following

day the membranes were washed four times in washing buffer for 15 min each prior to incubating with enzyme-conjugated secondary antibody for 2 hours at room temperature. Membranes were washed as before and 400 µL of Pierce Supersignal® West Dura Extended Duration Substrate (#34075, 1:1 ratio Luminol/Enhancer solution) were added to each membrane with protein side up. Membranes were placed protein side up in a cassette with a light activated Stratagene molecular weight marker and allowed to incubate in the dark for 5 min prior to film exposure. Kodak BioMax MR film (#Z350400, Sigma-Aldrich, St. Louis, MO, USA) was exposed for 2 to 10 min prior to development in a Kodak X-OMAT 2000A X-processor.

2.6. Flow cytometric analysis and immunofluorescence microscopy

Flow cytometry was performed using selected monoclonal antibodies on living *T. congolense* IL 3000 PCF and EMF and *T. simiae* CP-11 and *T. b brucei* 427 PCF. Trypanosomes were counted and suspensions (1×10^8 cell/mL) were prepared in complete PCF medium and kept on ice. To each reaction tube 50 µL cell suspension and 100 µL of relevant test mAb were added, followed by 20 min incubation on ice. To stop the incubation, one mL cold PCF medium was added and samples centrifuged at 800 × g for 1 min at 4 °C on a microcentrifuge. Supernatant was removed and the wash repeated once. Following this, 50 µL of fluorescent secondary antibody (1/50 dilution in PCF medium: Alexa Fluor® 488-conjugated goat anti-mouse IgG (H+L); A11001, Invitrogen, Eugene, Oregon, USA) and Alexa Fluor® 488-conjugated goat anti-mouse IgM (µ chain) (A21042, Invitrogen, Eugene, Oregon, USA) were added and cells incubated for 20 min on ice. Cells were washed as before and supernatant was fully aspirated. Cells were resuspended in 50

µL PCF medium. Seven hundred µL BD FACSFlow™ (342003, BD Biosciences, Mississauga, Ontario) were added to the remaining suspension. Cells were analyzed on a BD FacsCalibur flow cytometer. Controls included no primary antibody, mAb 491, and mAb 247.

Immunofluorescence microscopy was also used to examine monoclonal antibody binding to trypanosomes. Parasites from the same preparations used for flow cytometry were microscopically observed with a 100 x oil immersion objective using a fluorescence microscope (Zeiss Standard binocular microscope fitted with an epi-fluorescence attachment) by adding ten µL to a slide and covering with a #1 glass coverslip. Analysis of cell morphology, specifically the position of the kinetoplast and nucleus, was aided by using the DNA-binding dye 4',6-diamidino-2-phenylindole (DAPI).

Immunofluorescence microscopy was also performed on live *T. congolense* IL 3000 EMF adhered to coverslips, in addition to the suspensions mentioned above. Coverslips were incubated with *T. congolense* IL 3000 EMFs growing in a 24-well tissue culture plate. When the EMFs adhered and formed a carpet of cells on the coverslips, usually after two to three days, the coverslips were removed, washed with PBS to remove unbound trypanosomes (and potential MCFs) and the standard procedure for immunofluorescence flow cytometric analysis on live cells was followed.

Fluorescence microscopy was also performed on suspensions of fixed-permeablized procyclic and epimastigote culture forms of *T. congolense* IL 3000. Parasite samples were air-dried onto slides at room temperature for 30 min prior to acetone fixation at -20 °C for 20 min. Slides were then air dried to remove the acetone. Undiluted tissue culture supernatants containing mAbs of interest were incubated on the slides for 20 min at 37 °C

in a humidified chamber. Slides were rinsed with four quick washes by sequential immersion in three separate Coplin jars filled with PBS, pH 7.4. A 1:50 dilution of Alexa Fluor® 488-conjugated goat anti-mouse IgG (H+L) (A11001, Invitrogen, Eugene, Oregon, USA) and Alexa Fluor® 488-conjugate goat anti-mouse IgM (μ chain) (A21042, Invitrogen, Eugene, Oregon, USA) was added and the slides incubated for 20 min at 37 °C. Slides were washed as before and immunofluorescence was observed using a Leica DM6000B microscope fitted with an epifluorescence attachment and an oil-immersion 100 x Neofluor objective.

2.7. Infection of tsetse with T. congolense 1/148

In addition to my analysis of parasites grown *in vitro*, I also wanted to analyze the *in vivo* protein expression of parasites taken directly from tsetse. Since it is not legal to have live tsetse and trypanosomes in the same laboratory in Canada, a collaborator, Dr Lee Haines at the Liverpool School of Tropical Medicine, Liverpool, UK, agreed to supply parasites taken directly from infected tsetse. Male tsetse (*Glossina morsitans morsitans*) were infected with *T. congolense* 1/148 (Savannah type strain) (Young and Godfrey, 1983) by membrane feeding on 5 mL of defibrinated horse blood (TCS Biosciences Ltd., Buckingham, UK) containing 5 x 10^5 per mL *T. congolense* 1/148 BSF (Rosen *et al.,* 1981). The volume of blood consumed by tsetse fluctuates between 17 - 35 μL (depending on sex, degree of starvation and size of species), which corresponded to the ingestion of approximately 0.85 - 1.75 x 10^4 trypanosomes per fly. The flies were infected August 12, 2010 and either terminated early on August 23 (10 days post infection) in order to obtain early midgut forms, or late on September 6, 2010 (25 days post infection) in order to obtain

late midgut forms, proventricular (PV) forms and mouthpart forms (either epimastigotes or metacyclic forms). None of the flies dissected on August 23 showed mouthpart infections. To harvest trypanosomes, each fly was eviscerated into a drop of PBS on a slide and scored for infection. If positive, the PV and midgut were quickly severed, collected in separate microtubes containing 50 µl of PBS, vortexed and the material vigorously pipetted to liberate the trypanosomes from the tissue. The tubes were kept on ice for 30 min to allow the tissues to settle so that cleaner trypanosome preparations could be achieved. *T. congolense* 1/148 BSF parasites were taken directly from rats with high BSF parasitemias. All tsetse samples were air-dried onto poly-L-lysine microscope slides, acetone-fixed and probed with the relevant antibody as described in section 2.6.

Chapter 3. *T. congolense* Heptapetide repeat protein (TcHRP)

3.1. Introduction

In 2002, using radioactive GPI precursors, the Bütikofer lab discovered a trypsin-resistant 58 kDa protein present on the surface of *T. congolense* Kilifi procyclic culture forms (Bütikofer *et al.,* 2002). In their analysis of surface protein expression during fly transmission, the Bütikofer lab found that like *T. brucei* spp*., T. congolense* also expresses different GPI-anchored molecules that are also developmentally regulated during fly transmission. In tsetse, while *T. brucei* spp. express variable amounts of EP and GPEET procyclins (Bütikofer *et al.,* 1997; Treumann *et al.,* 1997; Ruepp *et al.,* 1997; Roditi *et al.,* 1998) with EP predominating after the seventh day (Vassella *et al.,* 2000; Acosta-Serrano *et al.,* 2001), *T. congolense* procyclic forms were frequently found to be negative for both PRS and GARP fourteen days post-infection, suggesting that another stage-specific surface molecule was expressed at this point in the life cycle (Bütikofer *et al.,* 2002). The Bütikofer lab in 2006 investigated whether the 50-58 kDa GPI-anchored protein might constitute the coat of these intermediate-stage parasites. Using a metabolic labeling approach similar to that used in their 2002 study, the Bütikofer lab partially purified the protein from *T. congolense* Kilifi procyclic forms. They obtained an N-terminal amino acid sequence, identified the encoding gene and generated a polyclonal antiserum which they subsequently used to analyze parasites taken directly from tsetse (Utz *et al.,* 2006). Comprehensive analyses revealed a mature protein consisting almost exclusively of 13 identical EPGENGT heptapeptides repeats; they named this molecule the *T. congolense* heptapeptide repeat protein (TcHRP) (Figure 3.1).

1	MIP
4	RQLCLLTVLLFGAALFGAERSGADASAPVVRRKVEVPLAKAD
46	**EPGENGT**EPGENGT**EPGENGT**EPGENGT**EPGENGT**EPGENGT
88	**EPGENGT**EPGENGT**EPGENGT**EPGENGT**EPGENGT**EPGENGT
130	**EPGENGT**EPAAGSVGFTSVVIPMAFAAFSMVATF

Figure 3.1. Deduced amino acid sequence of the TcHRP mRNA, a 58 kDa protein from T. congolense Kilifi procyclin Kil1 submitted November 2004. Accession number AY827559.

A *predicted N-terminal signal sequence and C-terminal GPI addition signal are underlined. The 13 identical EPGENGT heptapeptide repeats are indicated by alternating boldface type and lightface type (positions 46 – 136). (Modified from* Utz *et al.,* 2006*; with permission from Dr. P. Bütikofer).*

TcHRP undergoes extensive N-glycosylation on the polypeptide chain, with up to 13 high-Man-type oligosaccharides ranging from $Man_5GlcNAc_2$ to $Man_9GlcNAc_2$, as well as O-glycosylation linked via phosphodiester bonds to threonine (Thr) residues in the repeats. Taken together, TcHRP has a much higher apparent molecular mass than predicted based on the amino acid sequence alone. The lipid portion of the GPI anchor consists exclusively of *sn*-1-stearoyl-2-*lyso*-glycerol-3-HPO_4-1-(2-*O*-acyl)-D-*myo*-inositol and also has potential complex GPI modifications (Utz *et al.,* 2006). Similar proteins were identified in *T. congolense* Savannah procyclic forms. TcHRP is an abundant, novel GPI-anchored protein in *T. congolense* procyclic forms that shares a number of characteristics with *T. brucei* EP and GPEET such as being small, highly acidic and consisting almost exclusively of repetitive peptide sequences, and thus has become known as *T. congolense* procyclin. However, unlike the *T. brucei* procyclins which contain, at most, a single homogenous $Man_5GlcNAc_2$ glycan (Acosta-Serrano *et al.,* 1999), the *T. congolense* procyclins are heavily glycosylated

with at least 10 oligomannose N-glycans, possible phosphodiester-linked glycans and a GPI anchor with potential complex GPI modifications.

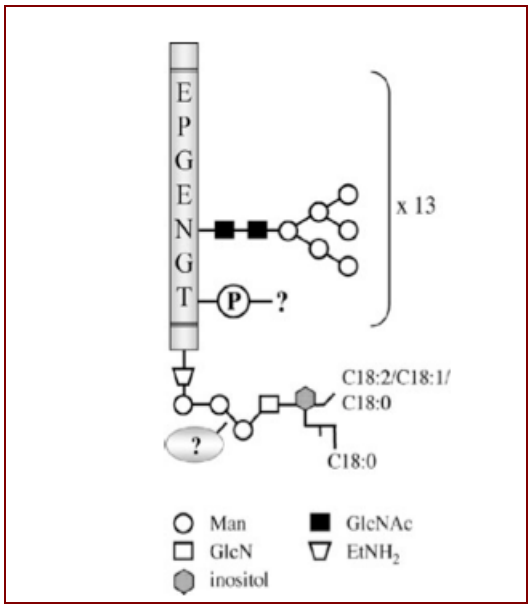

Figure 3.2. Schematic representation of the partial chemical structure of T. congolense Kilifi procyclins.
Shown are the proposed N-glycan structures and possible phophoglycan modifications linked to the EPGENGT repeats of the mature protein (Ala^{44}-Gly^{141}; gray bar) and the partial structure of the GPI anchor with the identified fatty acyl chains. (Adapted from UTz et al., 2006*; with permission from Dr. P. Bütikofer).*

TcHRP is interesting because it represents an intermediate-stage marker expressed by established procyclic forms in the fly midgut. TcHRP is expressed continuously during the course of a midgut infection and its N-terminal domain is shorter than predicted. In this respect, *T. congolense* procyclin (TcHRP) resembles *T. brucei* EP, which is expressed throughout the course of a midgut infection. During infection in the tsetse midgut, the N-terminal domains of EP and GPEET are quantitatively removed by proteolysis, leaving little more than the protease-resistant amino acid repeats (Acosta-Serrano *et al.,* 2001). These

repeat domains could form a shield that safeguards susceptible proteins on the cell surface from powerful gut proteases.

The lack of specific probes has made the characterization of specific surface proteins impossible. The first anti-*T. congolense* monoclonal antibody, mAb 491, made in the Pearson lab (Beecroft *et al.,* 1993) has been the only molecular probe available for *T. congolense* surface molecules. Unfortunately this mAb recognizes a carbohydrate epitope common to several *T. congolense* surface proteins: PRS, GARP, CESP and TcHRP. It has not therefore been determined which surface molecule appears first during parasite differentiation in the insect host. I set out to make a specific probe for TcHRP and to use it to study its differential expression during fly transmission. I did this in two different ways: by immunizing mice with either a purified butanol-extracted preparation of TcHRP or with *T. congolense* IL 3000 PCF lysates. The former was unsuccessful due to insufficient amounts of TcHRP present in the sample; the later was successful and will be discussed in detail below.

3.2. Materials and Methods

3.2.1. Trypanosomes

PCF of *T. congolense* IL 3000 (Fish *et al.,* 1989), *T. congolense* K45/1 (derived from *T. congolense* STIB744; Swiss Tropical Institute, Basel, Switzerland), *T. simiae* CP-11 (Zweygarth *et al.,* 1987; Majiwa *et al.,* 1987) and *T. b. brucei* 427-01 (Cross and Manning, 1973) were used.

3.2.2. Derivation of monoclonal antibodies

Two female BALB/c mice, aged six to eight weeks, were immunized with *T. congolense* IL 3000 PCF lysates. Monoclonal antibodies were generated according to Chapter 2; section 2.3.

3.2.3. Screening of monoclonal antibodies on glutaraldehyde-fixed trypanosomes

Hybridoma tissue culture supernatants were screened on glutaraldehyde-fixed *T. congolense* IL 3000 as an initial screen for surface reactivity. Trypanosomes were washed twice in ice-cold PBS pH 7.4 supplemented with 1 % glucose, followed by centrifugation at 850 × g for 10 min. Cell pellets were then resuspended to 2×10^7 cells/mL in PBS/ 1 % glucose, and 50 µL (1×10^6 cells/well) of cell suspension was added to each well of a BD Falcon™ flexible, U-bottom polyvinyl chloride RIA microplate (#353911, Franklin Lakes, NJ, USA). The plates were centrifuged at 50 × g for 5 min at 4 ^{o}C to settle the cells. Cells were fixed by adding 100 µL 0.25 % freshly prepared glutaraldehyde in PBS and incubated for 15 min at room temperature. The plates were centrifuged as before and the supernatant removed by inverting the plate and gently flicking to remove excess liquid. Plates were washed twice by adding 100 µL PBS. Normal ELISA protocol was followed after this (refer to Chapter 2, section 2.4). A selection of monoclonal antibodies that bound either the *T. congolense* IL 3000 lysates, glutaraldehyde-fixed *T. congolense* IL 3000 or both were chosen for further analysis.

3.2.4. Generation of EPGENGT peptide and screening of monoclonal antibodies

The sequence of the TcHRP repeat peptide portion, EPGENGT, was sent to JPT (Peptide Technologies GmbH, Volmerstrasse 5 (UTZ), 12489 Berlin, Germany) for custom peptide synthesis. They sent us 2 mg of lyophilized material, > 90 % HPLC purity, which was used to screen the anti-*T. congolense* IL 3000 mAbs in ELISA.

3.3. Results

3.3.1. Selection and analysis of monoclonal antibodies produced by immunization with T. congolense IL3000 procyclic culture forms

More than 400 mAbs obtained after immunization with trypanosome lysates were screened first by ELISA on lysates and then by a modified ELISA using glutaraldehyde-fixed *T. congolense* IL 3000 as solid-phase adsorbed antigen. A total of 96 stable clones were positive in these assays. Seventeen were positive on lysates only, 48 were positive on glutaraldehyde-fixed cells only and 31 were positive on both cell lysates and fixed cells. Since there are some contraindications to using glutaraldehyde-fixed cells as a test for surface reactivity (cell membrane integrity may be altered, allowing antibodies to bind to non-surface epitopes), positive results from this assay were viewed with caution. Surface reactivity of the mAbs was thus tested by immunofluorescence microscopy on live PCF. Of the 96 mAbs tested, 19 were reactive with the surface of living PCF, 13 of which had also tested positive on glutaraldehyde-fixed cells.

Immunoblot analysis of all 96 mAbs on SDS-PAGE separated proteins in lysates identified 25 reactive mAbs. The pool of 96 mAbs was then narrowed down to 28 mAbs which I considered a good representation of the initial 96: 16 worked on immunoblots only, 3 worked in immunofluorescence only, and 9 worked in both assays. Having a mAb that is

either surface- or immunoblot- reactive is useful but has its limitations: one is (usually) reactive to a topographically assembled epitope that is destroyed under denaturing conditions, while the other is (usually) reactive to a linear epitope that is otherwise concealed under non-denaturing conditions. Both kinds of mAbs are useful, but having a mAb that has the benefit of surface and immunoblot reactivity is incredibly useful, and for these reasons all 28 mAbs were chosen for further analysis in an attempt to identify the antigens that they recognized.

3.3.2. Identification of the epitope recognized by anti-TcHRP mAb 4-E5.

One of the major surface proteins of *T. congolense* procyclic forms is the procyclin TcHRP. This protein was not available to us either as a purified natural protein, nor as recombinant protein. However, a major portion of the molecule consists of the repeat sequence EPGENGT, making this a possible epitope. I therefore had this peptide synthesized and used it as a solid-phase antigen in a "peptide ELISA" for screening of my mAbs. One mAb, 4-E5, bound the EPGENGT sequence strongly, a pleasant surprise since immunizations to derive mAbs had been carried out using whole parasite lysates, which would contain many proteins with an abundance of PTMs that may or may not conceal the polypeptide chain. In addition, it is known that TcHRP is one of the most densely glycosylated parasite surface molecules ever reported. To obtain a mAb that recognized this peptide epitope was lucky indeed. Several initial attempts to obtain anti-TcHRP mAbs by immunizing with TcHRP-enriched material obtained from collaborators in Bern, Switzerland, Dr. Peter Bütikofer and Eva Greganova, had been unsuccessful, we believe because of the low amounts of material supplied, since the immunized mice produced very poor antibody titres.

3.3.3. Life cycle stage and species specificity analysis of anti-TcHRP mAb 4-E5 by flow cytometry, immunoblotting and immunofluorescence microscopy.

Flow cytometry was used to determine the surface reactivity of mAb 4-E5 on live PCF and EMF of *T. congolense* IL 3000. The anti-*T. congolense* IL 3000 test bleed, the anti-*. Nannomonas*-CHO mAb 491 and the anti-*T. b. brucei* mAb 247 were used as positive and negative controls respectively. MAb 4-E5 did not react with the surface of live PCF and EMF (Figure 3.3). It is possible that this mAb binds the heptapeptide repeat epitope on the TcHRP that is concealed under non-denaturing conditions, perhaps by a glycosylation or concealment within the tertiary structure of the TcHRP molecule.

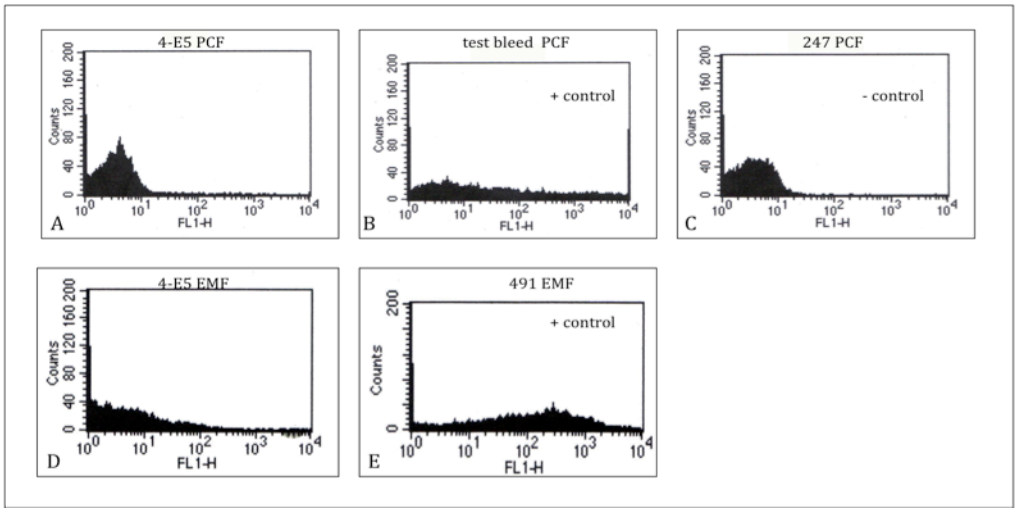

Figure 3.3. Flow cytometetric analysis of anti-TcHRP mAb 4-E5 on live PCF and EMF of T. congolense IL 3000.

PCF of T. congolense IL 3000 were incubated with either anti-HRP mAb 4-E5 (panel A), a 1/100 dilution of anti-T. congolense IL 3000 test bleed (panel B, positive control) or anti-T. b. brucei 427 EP repeat mAb 247 (panel C, negative control). EMF of T. congolense IL 3000 were incubated with either anti-HRP mAb 4-E5 (panel D) or anti-Nannomonas CHO mAb 491 (E, positive control). Undiluted tissue culture supernatants were used, followed by a 1:50 dilution of fluorescein-conjugated goat anti-mouse IgG/IgM secondary antibody. For every sample, ~ 5×10^6 cells were analyzed.

Analysis of the life cycle stage and species specificity for mAb mAb 4-E5 was then performed by immunoblot on SDS-PAGE separated proteins from lysates of PCF and EMF of

T. congolense IL 3000 grown *in vitro* and from lysates of PCF of *T. simiae* CP11 and *T. b. brucei* 427 (Figure 3.4). This assay expands on the flow cytometric results, by determining whether the linear epitope recognized by mAb 4-E5 was concealed under non-denaturing conditions by PTMs present on the surface of *T. congolense* IL 3000 EMF or whether it simply was not present. This mAb reacted with both the *T. congolense* IL 3000 PCF and EMF (Figure 3.4; lane 2 and 3), indicating it not to be life cycle stage-specific; intriguing, since TcHRP was reported by the Bütikofer lab to be predominantly expressed on the surface of intermediate established midgut forms. However, on closer inspection of the Utz paper, they do state "the occasional parasite isolated from the proboscis of a fly infected with *T. congolense* savannah also stained positive for TcHRP" (results were obtained using their polyclonal antiserum raised against purified natural HRP-data not shown). My mAb also reacted with the *T. b. brucei* and *T. simiae* CP 11 lysates, indicating it not to be species- or subspecies specific either (Figure 3.4; lane 4 and 5), a finding not totally unexpected, since the heptapeptide unit, EPGENGT, shares four amino acids, EPGT, in common with EP and GPEET repeats in *T. brucei*, possibly explaining the cross-reactivity of this mAb to protein bands in the *T. b. brucei* lysates (Figure 3.4; lane 5). An NCBI-BLAST search of the EPGENGT repeat unit revealed this sequence to be present in hypothetical proteins belonging to *Trypanosoma cruzi* strain CL Brenner, *Leishmania major* strain Friedlin, *Trypanosoma brucei gambiense* DAL972, and *Trypanosoma brucei* TREU927, among others. Given that no one has yet sequenced the genome of *T. simiae* it was not unexpected that it would show no hits for this short sequence. This mAb also reacted to multiple bands in all lysates tested, implying, at least for *T. congolense*, that it reacts to TcHRP having a range of PTMs. This would be consistent with TcHRP being modified with up to 13 high-mannose

oligosaccharides ranging from $Man_5GlcNAc_2$ to $Man_9GlcNAc_2$ as previously reported. In addition, the Thr residues in the EPGENGT repeats may also be modified with oligosaccharide chains linked via phosphodiester bonds (Utz *et al.,* 2006). Other possible explanations include variation between isolates/subspecies of *T. congolense*, or altered expression levels of TcHRP in different strains or culture media. The Bütikofer lab used *T. congolense* Kilifi STIB745 and *T. congolense* savannah TREU1457 (provided by the Swiss Tropical Institute, Basel, Switzerland) for their identification and analysis of HRP (Utz *et al.,* 2006), whereas I used *T. congolense* IL 3000 (another savannah strain).

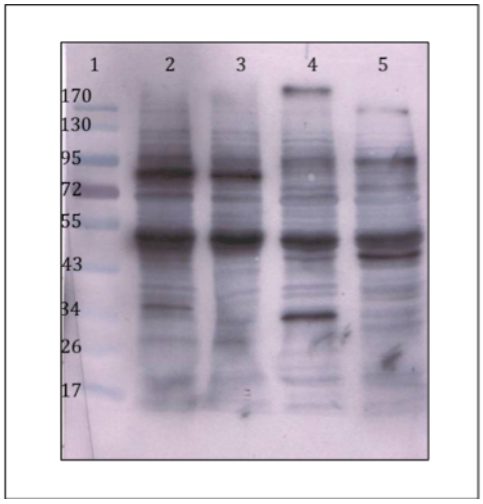

Figure 3.4. Immunoblot analysis of anti-TcHRP mAb 4-E5 on proteins in lysates of PCF and EMF of T. congolense IL3000 and PCF of T. simiae CP 11 and T. b. brucei 427, separated by 10% 1-D SDS-PAGE.

Shown are the nigrosin-stained PVDF membranes with immunoblot film overlaid. Lane 1 contained pre-stained MW markers (kDa), lanes 2, 3, 4 and 5 contained the lysates from 2×10^6 T. congolense IL 3000 PCF, T. congolense IL 3000 EMF, T. simiae CP 11 PCF, and T. b. brucei 427 PCF, respectively.

To evaluate if the EPGENGT epitope for mAb 4-E5 became exposed after acetone fixation/permeabilization and that the trypanosomes were of the correct life cycle stage, immunofluorescence analysis was performed on acetone-fixed PCF and EMF *T. congolense*

IL 3000 trypanosomes grown *in vitro* (Figure 3.5). Typically, in procyclic forms the nucleus is located centrally, with the kinetoplast posterior to the nucleus (Figure 3.5; top panel). As trypanosomes differentiate into epimastigote forms they become longer and the kinetoplast migrates anterior to the nucleus, towards the flagellum (Figure 3.5; bottom panel). The DNA-binding dye DAPI was used to aid in confirming the life cycle stage morphology of the procyclic and epimastigote culture forms used.

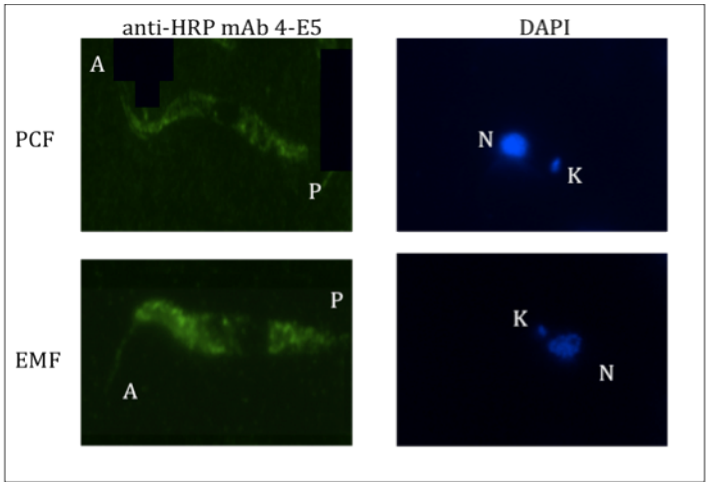

Figure 3.5. Immunofluorescence and DAPI analysis of anti-TcHRP mAb 4-E5 on acetone-fixed T. congolense IL 3000 PCF and EMF grown in vitro.
Top horizontal panel: acetone-fixed PCF. Bottom horizontal panel: acetone-fixed EMF. Undiluted tissue culture supernatant containing primary antibody was used, followed by a 1:50 dilution of Alexafluor 488 conjugated goat anti-mouse IgG/M secondary antibody. The kinetoplast (K) and nucleus (N) are indicted and the posterior (P) and anterior (A) ends of the trypanosomes have been labeled. Counterintuitively, the flagellum points to the anterior end of the trypanosomes.

Although anti-TcHRP mAb 4-E5 did not react with the surface of live PCF and EMF of *T. congolense,* it did react weakly with antigen in acetone-fixed/permeabilized PCF and EMFs of *T. congolense* suggesting that the epitope may be concealed by its position or arrangement of TcHRP in the cell membrane. It appears that TcHRP, at least on

trypanosomes grown *in vitro*, is not expressed exclusively by PCF but also by EMF. It would be valuable to test mAb 4-E5 on trypanosomes taken directly from tsetse infected with BSF parasites, thereby mimicking more closely a natural infection. In addition, having a 3-D crystal structure of TcHRP would shed much light on the packing/ arrangement and structure/function of this molecule. We anticipate that this will be the next step to understanding TcHRP.

Table 3.1. Summarized data for anti-TcHRP mAb 4-E5.

Isotype	IgM
Epitope	EPGENGT[1]
Species/ strain[2] specificity	not specific[3]
Life cycle stage[4]	PCF and EMF
Flow cytometry[5]	Internal
Immunofluorescence Microscopy[6] PCF (grown *in vitro*) EMF (grown *in vitro)*	 +1 +1→ +2

[1]Determined in ELISA on the synthesized peptide.
[2]Determined by immunoblot analysis on parasite lysates.
[3]Reacts with proteins in PCF lysates of *T. congolense* IL3000 (Savannah), *T. simiae* CP11 and *T. b. brucei* 427.
[4]Determined by immunoblot analysis on PCF and EMF of *T. congolense* IL 3000 grown *in vitro.*
[5]Trypanosomes were intact PCF and EMF of *T. congolense* IL3000. Positive mAb binding
[6]Epifluorescence microscopy on acetone-fixed *T. congolense* IL 3000 PCF and EMF grown *in vitro.* Fluorescence intensity was graded as very strong (+4) to weak (+1) or negative (-).

3.4. Discussion

The surface molecule TcHRP found on *T. congolense* surface was discovered in 2002 (Bütikofer *et al.,* 2002), and partially characterized in 2006 (Utz *et al.,* 2006). It was found

to be one of the most densely glycosylated parasite surface molecules ever reported and became known as the *T. congolense* procyclin. TcHRP, along with GARP, CESP and the protease resistant surface glycoconjugate (PRS), is temporally expressed through the *T. congolense* life cycle, and as with the others, its differential expression on the surface of various life cycle stages of *T. congolense* has not been well studied in part due to a lack of access to tsetse colonies and the absence of a suitable antibody specific to the TcHRP molecule.. I thus set out to derive and select specific monoclonal antibodies that could be used to measure HRP expression and used these reagents to study the TcHRP molecule in two major life cycle stages of *T. congolense* grown *in vitro.*

Initially, I tried making monoclonal antibodies by immunizing mice with TcHRP partially purified from *T. congolense* Kilifi STIB 745 (Būtikofer *et al.,* 2002), sent to us by the Būtikofer lab at the University of Bern, Bern, Switzerland. However, this approach was unsuccessful due to lack of sufficient quantities of TcHRP in the samples (obtained on two different occasions)- little or no immune response was achieved. I subsequently designed a different shotgun strategy by immunizing mice with *T. congolense* IL 3000 PCF lysates. MAb 4-E5 was ultimately selected based on its recognition in immunoblots of a protein with the same apparent molecular weight as TcHRP and its specific reaction in ELISA to a synthetic peptide (EPGENGT) that comprised the repeat portion of the TcHRP polypeptide. This was a very fortunate outcome in view of the fact that not only did I immunize mice with *T. congolense* lysates containing myriad different proteins with a variety of PTMs, but TcHRP itself is very densely glycosylated.

Given that peptide epitopes recognized by antibodies are usually between 5-8 amino acids in length, and that mAb 4-E5 reacted to the synthesized EPGENGT peptide, I was confident that I had identified the relevant epitope and had no doubt that this mAb reacted specifically with TcHRP in *T. congolense*. I could not, however, be sure as to what protein it was reacting with in the lysates of *T. simiae* CP 11 and *T. b. brucei* 427 since the former species has not been sequenced and the latter species expresses two GPI-anchored molecules, GPEET and EP procyclin, which share four amino acids with the *T. congolense* repeats, EPGT, possibly explaining the cross-reactivity seen with this mAb. In addition, the results of a BLAST search of EPGENGT revealed a number of hypothetical proteins belonging to *T. b. gambiense* that had regions of sequence similarity.

Using immunofluorescence microscopy on trypanosomes taken from the midgut of infected flies at different time points during the course of an infection, the Bütikofer lab demonstrated that while PRS was abundantly expressed on early PF and GARP was abundantly expressed by late-stage PF, TcHRP was continuously and approximately equally expressed by all midgut forms (Utz *et al.,* 2006). These authors also disclosed that "the occasional parasite isolated from the proboscis of a fly infected with *T. congolense* Savannah also stained positive for TcHRP" (data not shown) (Utz *et al.,* 2006). My immunoblot analysis showed mAb 4-E5 to be reactive with PCF lysates, as well as EMF lysates of *T. congolense* IL 3000 grown *in vitro*, thereby indicating that TcHRP was not exclusively expressed by procyclic forms grown *in vitro*. Due to insufficient material, I was unable to perform immunofluorescence on air-dried, acetone-fixed/permeablized trypanosomes taken directly from tsetse at different time points during a *T. congolense* infection. This

would certainly be worth doing, as it would give us a better idea of what was happening in an *in situ* infection.

I was able to generate a monoclonal antibody and to determine that it reacts specifically to the EPGENGT repeat of TcHRP, thereby making it possible to identify this molecule exclusively in *T. congolense*. This mAb was used experimentally to show that TcHRP is present in multiple forms, perhaps reflecting varying PTMs. It would be interesting to attempt to isolate some of these TcHRP variants to ascertain whether they are expressed from different genes or are simply due to PTMs. So far, only a single DNA sequence in *T. congolense* Kilifi encoding a protein with 13 identical EPGENGT repeats has been found. However, two genes in *T. congolense* Savannah strains encoding proteins with 11 and 13 heptapeptides have been identified, indicating that at least two different forms with different polypeptide chain lengths could be expressed from the *T. congolense* Savannah strain (IL 3000) used in my work (Utz *et al.,* 2006).

Chapter 4. Congolense Epimastigote-Specific Protein (CESP)

4.1. Introduction

One of the most crucial stages in the transmission of trypanosome infections is the development of the metacyclic trypanosome in the tsetse vector. In *T. congolense,* this development takes place in the proboscis (mouthparts) of tsetse. Here, *T. congolense* epimastigote trypanosomes cling by hemidesmosomes formed between their flagella and the chitinous wall of the food canal in the tsetse proboscis. The adherent parasites transform into free-swimming, VSG-expressing, animal-infective trypomastigote metacyclic forms, ready for discharge with the saliva when the fly bites its next victim (Evans *et al.,* 1979; Gray *et al.,* 1981; Tetley *et al.,* 1985; Vickerman, 1969). Molyneux *et al.,* (1979) hypothesized that the presence of epimastigote clusters in the mouthparts may interfere with the electrophysiologic control of probing, making a fly with a proboscis infection probe much more frequently than a non-infected fly. It was clear even back then that epimastigote attachment was important in trypanosome transmission, but it was not clear if attachment was simply to provide a holdfast to retain epimastigotes in a region where the metacyclic progeny can infect the vertebrate host or whether it was important for cyclical development of the trypanosome. The answer to this came in 1988, when it was concluded that attachment has developmental significance. It is a necessary prerequisite for metacyclogenesis, but not epimastigote proliferation (Hendry *et al.,* 1988). Hendry's lab found that preventing *T. congolense* from attaching to culture flasks by shaking the cultures or by providing a polypropylene substratum did not inhibit epimastigote division, but did prevent the differentiation of epimastigotes into metacyclics.

Recently, a novel 100 kDa stage-specific GPI-anchored surface glycoprotein with potential adhesion activity and which was selectively expressed in the epimastigote life cycle stage of *T. congolense,* was discovered and named congolense epimastigote-specific protein (CESP) (Sakurai *et al.,* 2008). These researchers observed that the culture supernatants of EMF but not PCF promoted adhesion of PCF parasites in an *in vitro* assay. Using biosynthetic labeling of PCF and EMF proteins they identified a 100 kDa protein, expressed exclusively by EMF, that accumulated in the tissue culture supernatants. CESP is encoded by a multi-copy gene with an open reading frame of 2070 base pairs (bp) encoding a polypeptide of 689 amino acids. The translated sequence of CESP (Figure 4.1) revealed that the protein consists mainly of highly hydrophilic amino acids with short hydrophobic N- and C-terminal peptides, predicted to be a signal peptide with GPI-anchor binding region, respectively.

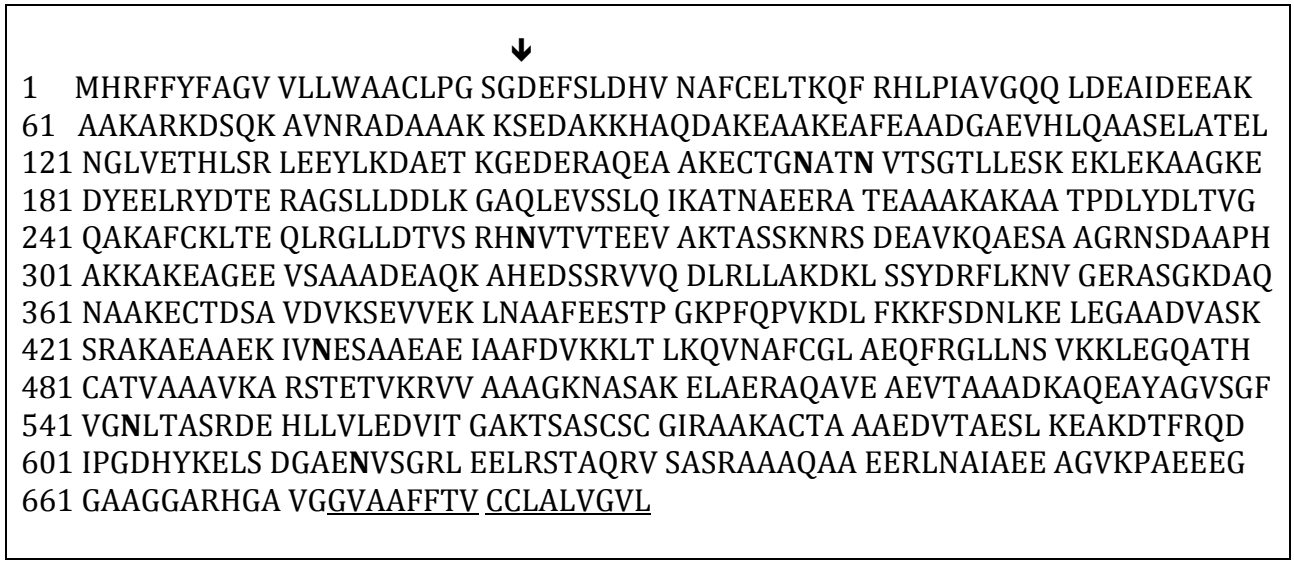

Figure 4.1. The protein sequence of congolense epimastigote-specific protein (CESP). *First published amino acid sequence of CESP (Sakurai, T., Sugimoto, C. and Inoue, N. in 2007*; *Accession number* BAF91109). *Six asparagine residues predicted to be* N*-glycosylated are bolded. The putative cleavage site of a predicted signal peptide is designated by an arrow. The C-terminal hydrophobic region is underlined.*

Using confocal laser scanning microscopy, the Inoue lab showed CESP to be specifically expressed on the EMF cell surface. Since CESP was initially identified in EMF tissue culture supernatants, it is likely that CESP is also secreted or released by dying parasites or some other mechanism. Much like all the other trypanosome-specific GPI-anchored molecules, CESP was found to have a higher apparent molecular mass than predicted from the translated amino acid sequence. This suggested that CESP was highly glycosylated. In fact, CESP has six predicted N-glycosylation sites. Similar to GARP, CESP was predicted to be highly α helical. This is noteworthy because like GARP, CESP is expressed at the juncture between expression of surface molecules on insect forms (in this case epimastigotes) and VSG expression on metacyclic forms. Thus, CESP may be a transition molecule that prepares the membrane for expression of VSG, much like the limited expression of GARP potentially mitigates the loss of VSG and eases the transition between the two disparate environments.

CESP is the first and only molecule found to date that is expressed exclusively by *T. congolense* EMF. Since attachment of epimastigotes is necessary for metacyclic production but not for epimastigote proliferation, the identification of CESP and its potential as an adhesion molecule that allows *T. congolense* epimastigotes to adhere to the tsetse proboscis is worth further investigation. I set out to make monoclonal antibodies to CESP for use in expression analysis throughout the trypanosome life cycle and for epitope mapping of CESP.

4.2. Materials and Methods

Cell fusion was performed using spleen cells from mice immunized with recombinant CESP and monoclonal antibodies were derived, screened, and analyzed according to outlined protocols described in Chapter 2 (sections 2.3-2.6).

4.2.1 Epitope mapping of selected monoclonal antibodies

To prepare rCESP for in-solution digestion with trypsin, 50 µg per reaction of recombinant protein was first concentrated overnight at 4 °C by precipitation using a 9:1 volume ratio of ice-cold acetone to protein. The protein precipitates were centrifuged at 10,000 rpm (7378 x g) for 1 min on a SIGMA 1-14 microcentrifuge (Osterode, Germany) to pellet the precipitate. Acetone was removed and the precipitate resolubilized in 30 µL 25 mM ammonium bicarbonate in 6 M urea. For analysis by mass spectrometry, samples were reduced with 30 µL 10 mM dithiothreitol (DTT) for 45 min at 37 °C, alkylated with 30 µL 40 mM iodoacetamide for 45 min at 37 °C in the dark. Reactions were quenched by addition of 30 µL 40 mM DTT. Following reduction and alkylation, rCESP was digested overnight at 37°C, using a trypsin to protein ratio of 1:20 (20 ng/µL stock, #V5111, sequencing grade modified porcine trypsin, Promega, Madison, WI, USA). To stop digestion 100x protease inhibitor cocktail set V, EDTA-free (#539137, Calbiochem®, La Jolla, CA, USA) was added to a final concentration of 1x. The rCESP digest was diluted with PBS to give a final concentration of 0.2 µg CESP/ µL.

Immunocapture mass spectrometry was used to try and identify the peptide epitopes bound by monoclonal antibodies. The murine monoclonal antibodies were first captured on magnetic beads. These solid-phase adsorbents were used to capture peptide

epitopes from the rCESP enzyme digests. The beads chosen for capturing the monoclonal antibodies of interest were goat anti-mouse IgG dynabeads (#110.33, Invitrogen Dynal AS, Oslo, Norway). The desired volume of Dynabead slurry (25 µL of Dynabeads, $\sim 1 \times 10^7$ beads) was transferred to a microfuge tube, followed by three quick washes in 500 µL PBS each. The tubes were placed on a magnet (#2010, Dynal® Invitrogen bead separator) for 1 min and the supernatant discarded. According to the manufacturer's instructions, 25 µL of Dynabeads can capture 0.5 – 1.5 µg of primary antibody. Therefore each 25 µL of beads were incubated with 1 mL of specific monoclonal antibody tissue culture supernatant (which contains approximately 1 – 10 µg/mL specific antibody) overnight at 4 °C on a rotator. An excess of mAb was used to try and ensure complete bead saturation. The next day the tubes were placed on a magnet and the mAb supernatant removed. The beads were washed three times with 1 mL PBS. After the final wash, the beads were resuspended in PBS to obtain the original volume, 25 µL, and stored at 4 °C.

For each peptide capture reaction 50 µL of digested rCESP (0.2 µg/µL, ~10 µg) were added to 25 µL of mAb-coated Dynabeads. Capture reaction mixtures were incubated at room temperature for 1.5 hours with shaking. Following peptide capture, the tubes were placed on a magnet and the supernatant removed and saved (unbound, non-epitope peptides). The beads were washed with three times with 1 mL PBS to remove unbound or weakly bound peptides. To elute the bound peptides from their capture antibodies, 50 µL of 5 % acetic acid were added to the beads and the mixture was incubated for 5 min at room temperature. Finally, the supernatant was removed and saved for mass spectrometric identification of peptide epitopes.

4.2.2. MALDI-TOF mass spectrometry

Antibody-captured peptides, non-captured peptides from trypsin digestion, and peptides from complete trypsin digest (control) were desalted and concentrated using ZipTips (C18 resin; P10, Millipore Corporation, Bedford, MA, USA). For each sample, 1 µL of the desalted, concentrated peptide mixture was mixed 1:1 with the matrix, alpha-cyano-4-hydroxycinnamic acid and spotted onto a Voyager, 100 position, stainless steel MALDI plate (Applied Biosystems, Foster City, CA, USA). 'Calmix 2' internal standards (purchased from UVic mass spec. facility), angiotensin 1 (1296.6853), ACTH 1-17 clip (2093.0867) and ACTH 18-39 clip (2465.1989) were used to calibrate the mass spectrometer prior to analysis of samples. An Applied Biosystems Voyager DE-STR mass spectrometer (Applied Biosystems, Foster City, CA, USA) running in delayed extraction, reflectron mode was used to acquire MALDI-TOF data. The CESP protein sequence was submitted for *in silico* digestion using MS-digest (Protein Prospector software package; San Francisco, CA, USA: http://prospector.ucsf.edu/) and peptide masses were compared to the in-solution digests.

4.3. Results

4.3.1. The selection and analysis of anti-CESP antibodies mAb 1-D11 and 3-C6 by ELISA, immunoblot, flow cytometry and immunofluorescence.

Screening hybridoma supernatants by ELISA and immunoblotting on recombinant CESP-GST fusion protein resulted in the selection of 12 stable clones that were chosen for further analysis. Only two mAbs, 1-D11 and 3-C6, were found to react with rCESP-GST in both assays. Immunoblot analyses were performed using lysates of *in vitro* grown *T.*

congolense IL 3000 (savannah strain) PCF and EMF and of PCF of *T. congolense* K45/1 (Kilifi strain), *T. b. brucei* 427 and *T. simiae* CP11. The results for mAb 3-C6 are shown in Figure 4.2. Immunoblot results for mAb 1-D11 were identical (data not shown). The results indicate that both mAbs recognize linear epitopes on the native CESP in *T. congolense* IL 3000 EMFs exclusively, confirming that they are specific to the EMF lifecycle stage. This is in agreement with the finding that *CESP* mRNA is exclusively expressed in the EMF stage (Sakurai *et al.,* 2008). Neither mAb reacted to any other trypanosome lysates tested, confirming their specificity for *T. congolense* IL 3000 EMF. This was somewhat expected since none of the other trypanosome species transformed into EMFs, even after extended culturing. Flow cytometric analysis was used to determine the surface reactivity of mAbs 1-D11 and 3-C6 on live PCF and EMF of *T. congolense* IL 3000 (Figure 4.3). My analysis also included live PCF of *T. simiae* CP11 and *T. b. brucei* 427 (results not shown). Only mAb 3-C6 displayed surface-reactivity, exclusively on EMFs, indicating that its epitope is linear, unobscured by PTMs such as glycosylations, and surface exposed. In view of the fact that these parasites are so heavily glycosylated, this is significant. On the other hand, 1-D11 was not surface-reactive, implying that the linear epitope was concealed by a glycosylation or shielded by its position within the tertiary structure under non-denaturing conditions.

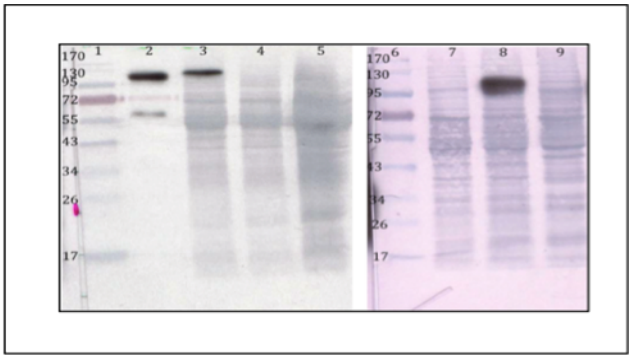

Figure 4.2. Immunoblot analysis of anti-CESP mAb 3-C6 on proteins from lysates of in vitro grown trypanosomes: T. congolense IL 3000 EMF and PCF, and PCF of T. congolense K45/1, T. simiae CP11 and T. b. brucei 427.

Shown are the nigrosin-stained PVDF membranes with immunoblot film overlaid. The primary antibody used was mAb 3-C6. Lanes 1 and 6 contained pre-stained MW markers (kDa), lane 2 contained 0.4 µg of rCESP-GST fusion protein, lanes 3 and 8, 4, 5, 7, and 9 contained lysates from 5×10^6 T. congolense IL 3000 EMF, T. congolense K45/1 PCF, T. simiae CP-11 PCF, T. congolense IL 3000 PCF, and T. b. brucei 427 PCF, respectively.

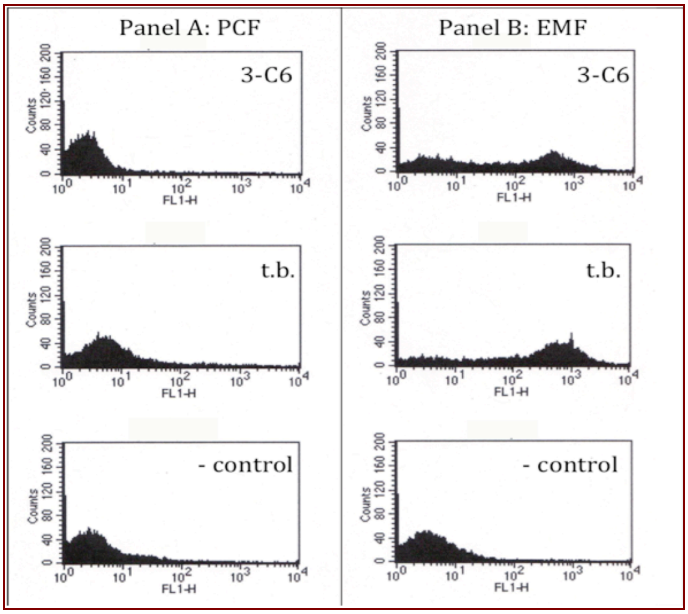

Figure 4.3. Flow cytometric analysis of anti-CESP mAb 3-C6 on live T. congolense IL 3000 PCF and EMF grown in vitro.

Panel A: PCF. Panel B: EMF.
Cells were incubated with undiluted mAb 3-C6 tissue culture supernatant (top row), a 1/100 dilution of anti-CESP test bleed (middle row; positive control) or no antibody (bottom row; negative control). The secondary antibody was fluorescein-conjugated goat anti-mouse IgG/IgM used at 1:50 dilution. For every sample, ~ 5×10^6 cells were analyzed.

Immunofluorescence microscopy using mAb 3-C6 was performed on *in vitro* grown, live EMF adhered to coverslips to evaluate the surface coverage and distribution of CESP on adherent parasites (Figure 4.4). The immunofluorescence revealed CESP to be expressed over the entire surface of the parasites with the greatest intensity occurring at the anterior position where the flagellum attaches to the coverslip (Figure 4.4; Panel A and B). This would be in accordance with confocal laser scanning microscopy studies done by the Inoue lab in Obihiro, Japan, in which they established, using polyclonal antisera, that CESP was expressed and localized on the EMF surface (Sakurai *et al.,* 2008). The high intensity of fluorescence at the attachment site also adds, albeit indirectly, to the evidence that CESP may be an adhesion molecule that is required by EMF for adherence to the tsetse mouthparts and subsequent metacyclogenesis. Additionally, the Inoue lab found that their mouse polyclonal antisera showed weak activity with the flagellum of BSFs. This is in keeping with our hypothesis that CESP, like GARP, may also be coexpressed with VSG and may facilitate the parasite's transition from EMF to VSG-expressing MCF and BSF.

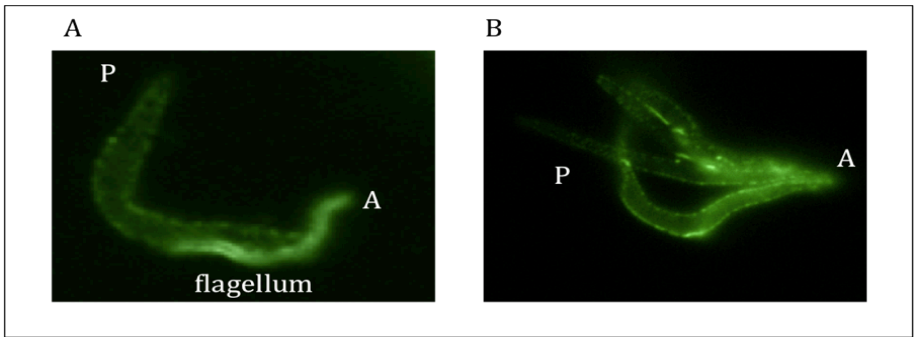

Figure 4.4. Immunofluorescence analysis of mAb 3-C6 on live, adherent T. congolense IL 3000 EMF grown in vitro.

Posterior (P) and anterior (A) positions are indicated. Panel A shows the increased expression of CESP along the flagellum. Panel B shows the anterior attachment of EMFs to the coverslip. Undiluted mAb 3-C6 tissue culture supernatant was used as the primary antibody, followed by a 1:50 dilution of Alexafluor 488 conjugated goat anti-mouse IgG/M secondary antibody.

To determine the surface reactivity of mAb 3-C6 with parasites taken directly from infected tsetse, immunofluorescence was performed on *T. congolense* 1/148 (savannah type strain) parasites (Young and Godfrey, 1983) harvested from the midgut, proventriculus (PV) and mouthparts of infected tsetse (Figure 4.5), as well as BSFs harvested from the blood of infected rats. The DNA-binding dye DAPI was used to confirm the life cycle stage of the cells.

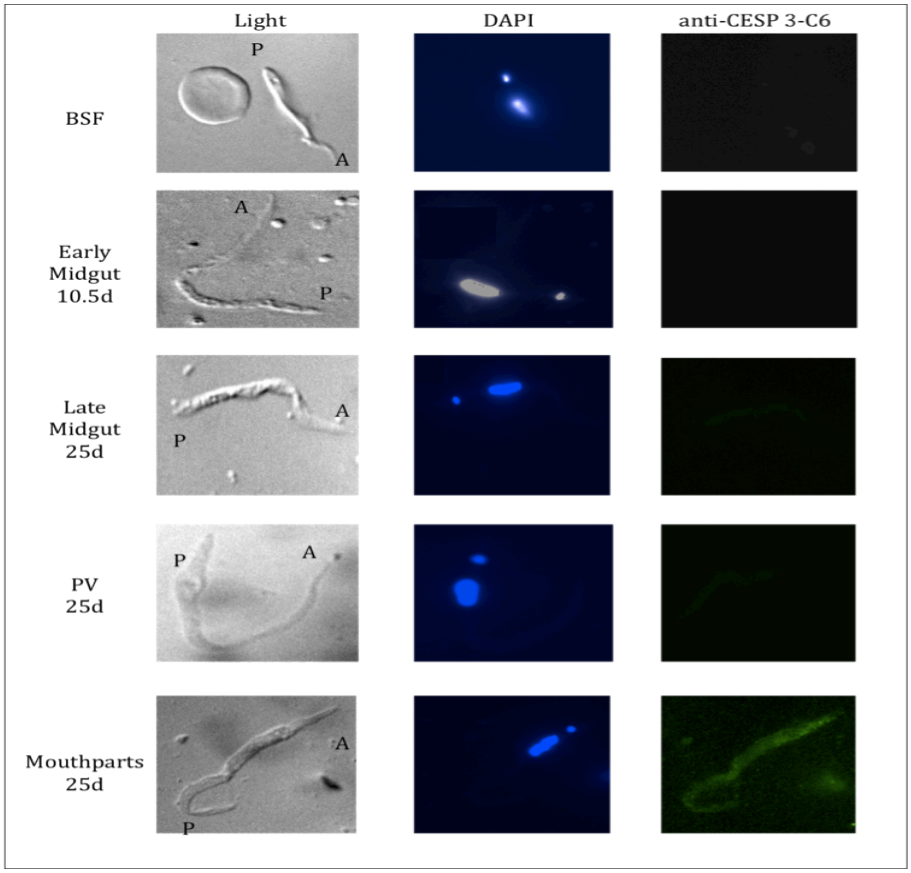

Figure 4.5. Immunofluorescence analysis of CESP expression in different life cycle stages of T. congolense taken directly from blood or from tsetse.

T. congolense 1/148 Savannah BSF isolated from infected rats (first panels); procyclic forms isolated from the midguts of infected tsetse after 10.5 and 25 days post infection (second and third panels); proventricular (PV) forms (fourth panel); epimastigote forms isolated from the mouthparts 25 days post infection(fifth panel). Trypanosomes were simultaneously labeled with anti-CESP 3-C6 and the DNA-binding dye DAPI to confirm their life cycle stage. The posterior (P) and anterior (A) ends have been indicated.

Unlike previous results reported from the Inoue lab, mAb 3-C6 did not show any immunofluorescence with BSF of *T. congolense*. This antibody also showed no fluorescence with any other life cycle stages except with those trypanosomes obtained from the tsetse mouthparts (Figure 4.5; fifth panel). Only these trypanosomes were found, based on the identification and localization of the kinetoplast using DAPI, to be true epimastigote forms. The immunofluorescence observed on these mouthpart forms was weak, especially when compared to the *in vitro* grown *T. congolense* IL 3000 EMF (Figure 4.4). It is possible that the different *T. congolense* strain used, 1/148 versus IL 3000, to infect the tsetse was responsible for the disparate immunofluorescence results observed for this mAb. Although being often representative of *in vivo* systems, *in vitro* culture systems still suffer from a lack of true environmental factors that occur *in vivo* and thus may alter protein expression. Much like the PCF of *T. congolense* IL3000 used in the Pearson lab at UVic, which express GARP at an intermediate level compared to established midgut procyclic forms in the tsetse (which do not express GARP), the epimastigote culture forms appear to express CESP at a level higher than those EMF from infected flies.

My results are in agreement with the immunoblot and flow cytometric results presented in Figures 4.2 and 4.3, and with the finding that CESP was selectively expressed on epimastigote culture forms (Sakurai *et al.,* 2008). I had to use tsetse-derived trypanosomes fixed to microscope slides and shipped from the Liverpool School of Tropical Medicine, but it would have been more favorable to compare the immunofluorescence on live trypanosomes taken directly from the infected tsetse. Since *T. simiae* and *T. godfreyi* also belong to the subgenus *Nannomonas*, neither of which easily differentiate into EMF, it remains to be determined whether CESP is expressed by all species of this subgenus or

whether it is exclusive to *T. congolense*. A summary of the anti-CESP mAb characterization is shown in Table 4.1

Table 4.1. Summarized data: characterization of selected anti-CESP monoclonal antibodies.

	mAb 1-D11	**mAb 3-C6**
Isotype	IgG_2a	IgG_2a
Species/ strain[1] specificity	*T. congolense* specific	*T. congolense* specific[2]
Lifecycle stage[3]	EMF	Mainly EMF[4]
Flow cytometry[5]	Internal	Surface
Immunofluorescence microscopy[6] EMF	Not Done	 +4
Immunofluorescence microscopy[7] PCF (grown *in vitro*) EMF (grown *in vitro)*	Not Done	 - +3
Immunofluorescence microscopy[8] BSF (from rat blood) Early midgut (from tsetse) Late midgut (from tsetse PV (from tsetse) Mouthparts (from tsetse)	Not Done	 - - +1 +2 +3

[1]Immunoblot analysis on PCF parasite lysates.
[2]Reacts to proteins on surface of acetone-fixed *T. congolense* IL 3000 and 1/148 strains.
[3]Immunoblot analysis on PCF and EMF lysates of *T. congolense* IL 3000 grown *in vitro.*
[4]Epifluorescence microscopy on acetone-fixed *T. congolense* PCF and EMF grown *in vitro,* as well as on acetone-fixed *T. congolense* 1/148 taken from tsetse *in vivo.*
[5]Trypanosomes were intact PCF and EMF of *T. congolense* IL3000. Positive mAb binding was labeled as surface reactive.
[6]Performed on intact, live EMF self-adhered to slide covers. Fluorescence intensity was graded as very strong (+4) to weak (+1) or negative (-).
[7]Epifluorescence microscopy on acetone-fixed *T. congolense* IL 3000 PCF and EMF grown *in vitro.*
[8]Epifluorescence microscopy on *T. congolense* 1/148 trypanosomes (*from tsetse*) absorbed to poly-L-lysine coated slides.

4.3.2. Validation of expression, purification and cleavage of rCESP generated in the Boulanger lab by immunoblotting, and its use in in-solution digestion to determine anti-CESP mAb 1-D11 and mAb 3-C6 epitopes.

I thought it prudent to try to identify the relevant epitopes recognized by mAbs 1-D11 and 3-C6. Drs. Inoue and Sakurai provided a plasmid construct containing part of the CESP sequence for recombinant expression and the Boulanger lab were successful in expressing, purifying, cleaving the GST tag and crystallizing CESP. They gave to me purified rCESP to use in my epitope identification experiments.

The reactivity of mAb 3-C6 to, and the successful expression, purification and cleavage of this rCESP was confirmed by immunoblotting (Figure 4.6). MAb 3-C6 reacted with this rCESP at its expected mass of 73 kDa, as well as to additional protein bands at ~50, 69, and 100 kDa (Figure 4.6; lane 2). The lower molecular weight bands may correspond to precursor/ breakdown products, whereas the higher molecular weight band may be some uncut CESP-GST fusion protein. While mAb 3-C6 reacted to several bands in the recombinantly expressed GARP, it reacted to only a single band at ~ 100 kDa in the *T. congolense* IL 3000 EMF lysates. The location of this band at 100 kDa in the *T. congolense* IL 3000 EMF lysates is in agreement with the apparent MW of CESP with a GPI anchor and PTMs (Sakurai *et al.,* 2008). The 50 and 69 kDa bands could be precursor/breakdown products that are either obscured by a PTM in the parasites, or are artifacts of the recombinant expression system.

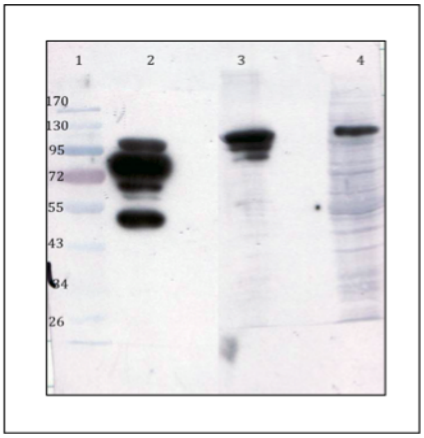

Figure 4.6. Immunoblot analysis and comparison of binding of anti-CESP mAb 3-C6 to rCESP produced at UVic or at Obihiro University.

Shown is the nigrosin-stained PVDF membranes with the immunoblot film overlaid. Lane 1 contained protein pre-stained MW markers (kDa), lane 2 and 3 contained 2 µg and 1 µg rCESP from the Boulanger and rCESP-GST from the Inoue lab, respectively, and lane 4 contained the lysates from 2 × 10^6 T. congolense IL 3000 EMF.

To attempt to determine the epitope recognized by mAbs 1-D11 and 3-C6, rCESP was precipitated, resolubilized, reduced, alkylated and cleaved with trypsin according to the protocol outlined in Chapter 4, section 4.2.1 and 4.2.2, to generate peptides for antibody-capture and analysis by MALDI-TOF mass spectrometry. Aside from a very good digestion of CESP with trypsin (Figure 4.7; Panel A), the results of this experiment were not positive. Neither mAb enriched any peptides (Figure 4.7; Panel B), indicating that trypsinization of CESP had destroyed the epitope. Or that the mAbs were of too low affinity to bind and hold peptides through the immunoenrichment procedure. Unfortunately, the difficulties in expressing, purifying and cleaving a recombinant protein precluded me from performing further digests of CESP with either GluC or α-chymotrypsin in an attempt to find the relevant epitopes.

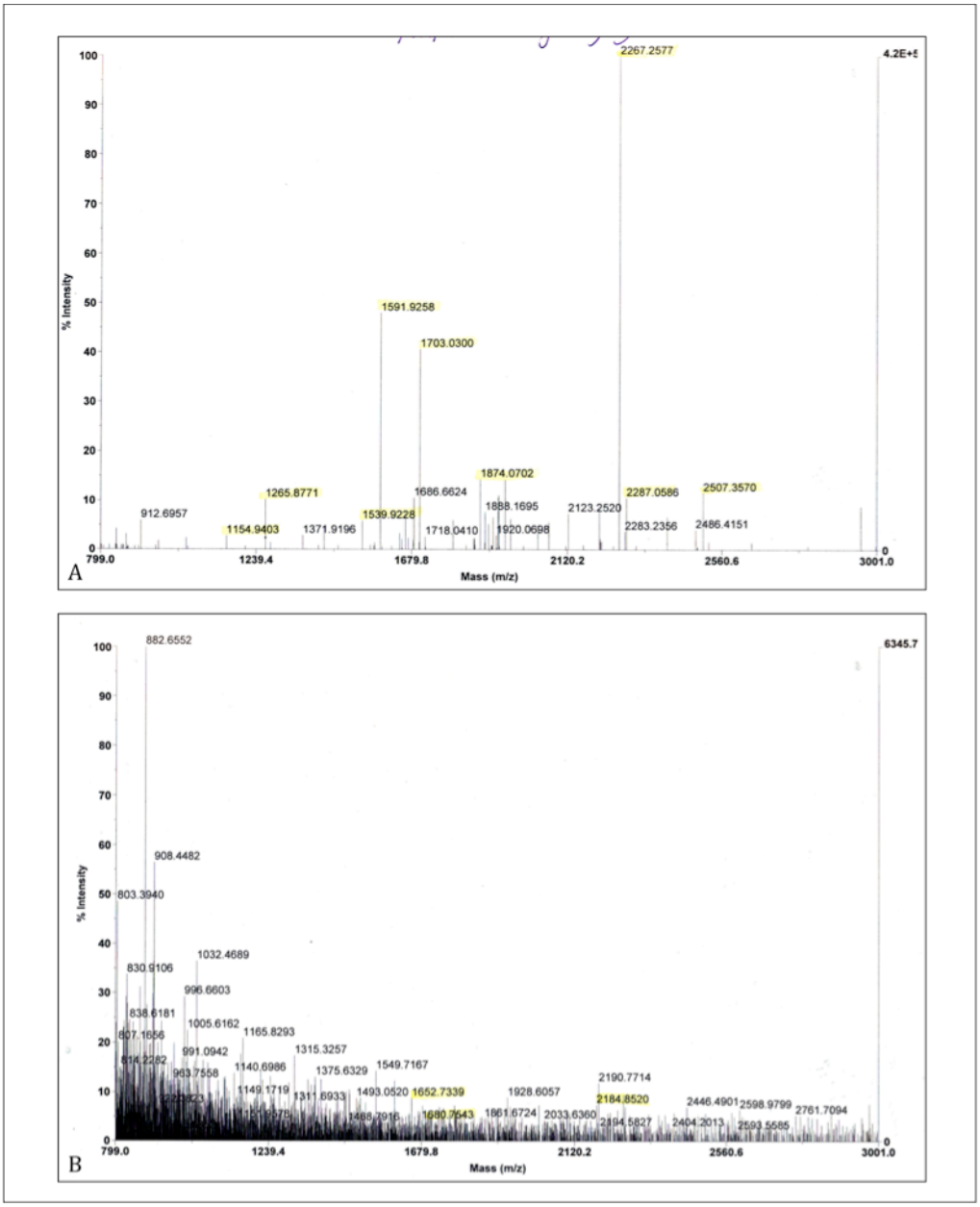

Figure 4.7. MALDI-TOF mass spectra of peptides from trypsin digested rCESP and epitope peptides captured by mAb3-C6.

Panel A: CESP trypsin digest (control). Panel B: peptides captured by mAb 3-C6. rGARP was digested with the serine protease trypsin to generate peptides that were then incubated with magnetic Dynabeads coated with mAb 3-C6 or mAb 1-D11 (spectrum not shown). Panel A shows that the digestion with trypsin was successful, with a variety of tryptic peaks corroborated by comparison with an in silico digestion of CESP. Panel B shows that no peptides were enriched. The same spectrum was observed when mAb 1-D11 was used for peptide capture (spectrum not shown).

4.4. Discussion

CESP is the most recently discovered *T. congolense* surface protein (Sakurai *et al.,* 2008). Its discovery and potential as an adhesion molecule that appears to allow *T. congolense* epimastigotes to adhere to the tsetse proboscis is very significant. It has been acknowledged as far back as 1979 that epimastigote attachment is important in trypanosome transmission, such that attachment is a necessary prerequisite for metacyclogenesis, but not epimastigote proliferation, in *T. congolense* (Hendry *et al.,* 1988). The Hendry's lab found that preventing *T. congolense* from attaching to culture flasks by shaking the cultures or by providing a polypropylene substratum did not inhibit epimastigote division, but did prevent the differentiation of epimastigotes into metacyclics.

The Noboru lab performed a similar adhesion assay in which they cultured PCF in the wells of culture plates that had been previously treated with epimastigote supernatant. They found that the PCF rapidly become firmly attached to the treated wells, but did not differentiate into EMFs. They isolated and partially characterized the unknown adhesion factor present in the EMF-supernatant and called it CESP. Could CESP be responsible for the adhesive property of EMFs and if so is it responsible for metacyclogenesis? As part of my thesis I generated monoclonal antibodies to CESP that can be used to analyze the expression of this molecule throughout trypanosome lifecycle, thereby shedding more light on the possible function of this adhesive molecule.

I derived two monoclonal antibodies specific for the CESP polypeptide. Having a mAb that is specific to the polypeptide enables researchers to analyze CESP without there being any cross-reactivity through PTMs, such as those that exist for mAb 491 which reacts to a carbohydrate moiety present on several *T. congolense* surface molecules, and

mentioned in detail throughout this thesis (Bütikofer *et al.,* 2002). One of these mAbs was reactive to the surface of living EMFs grown *in vitro* and acetone-fixed/permeabilized EMFs taken directly from infected tsetse mouthparts, as well as to lysates separated by gel and detected on immunoblot. This indicates that its epitope is both linear, not concealed by PTMs or its tertiary position on the protein. I would very much like to have been able to test this mAb on the surface of living parasites taken directly from the different tsetse compartments over the course of an infection.

As mentioned in the results section I was unable to determine the epitope for either anti-CESP mAb, due to lack of sufficient material to perform subsequent analyses with. I can, however, comment that whatever the epitopes may be they are destroyed after cleavage with trypsin.

A lot of study still needs to be done on CESP. It remains to be determined if this in indeed the adhesion molecule that is required by EMF metacyclogenesis into MCFs, or whether there is another as of yet undiscovered adhesion molecule that is responsible. Immunofluorescence analysis on living *T. congolense* EMF clearly showed CESP to be distributed over the entire EMF surface, with increased localization occurring along the flagellum. CESP concentration at the attachment site gives very good evidence that CESP is indeed an adhesion molecule. With my antibodies one could test the potential of CESP as an adhesion molecule required for metacyclogenesis both *in vitro,* and *in vivo* by engineering the trypanosome bacterial endosymbiont, *Wigglesworthia glossinidia*, to express single chain version of these mAbs, or incubating cultures in the presence of these mAbs and analyzing whether they prevent new EMFs from differentiating into MCFs.

Chapter 5. Glutamic Acid/Alanine-Rich Protein (GARP)

5.1. Introduction

The transformation of trypanosomes from BSF to procyclic forms is subject to a dramatic change in environment and therefore it is not surprising that their response, in terms of surface architecture, is equally dramatic. In *Trypanosoma* (*Trypanozoon*) *brucei* the synthesis of procyclin mRNA, and expression and incorporation of the GPI-anchored, invariant procyclin proteins into the surface membrane are the earliest known markers of this transformation event, and occur before the VSG coat is lost (Roditi *et al.,* 1989). In *T. brucei* procyclic and epimastigote forms, the surface coat consists of EP and GPEET glycoproteins, two classes of related proteins of the procyclin family which contain extensive tandem repeat units (Colmerauer *et al.,* 1989; Mowatt *et al.,* 1987; Roditi *et al.,* 1987; Roditi *et al.,* 1989). After the discovery of the EP and GPEET procyclins in *T. brucei*, two independent labs examined the surface of *T. congolense* using different approaches that ultimately converged, discovered the first cell surface molecule, a protein called glutamic acid/alanine-rich protein (GARP) (Beecroft *et al.,* 1993; Bayne *et al.,* 1993). The Bayne lab cloned the gene (Figure 5.1) by differentially screening a procyclic *T. congolense* cDNA library. The Beecroft lab used monoclonal antibody probes in conjunction with cation-exchange chromatography and gel electrophoresis to purify GARP and amino acid microanalysis and protein microsequencing to identify the sequence.

At the time GARP was called the *T. congolense* procyclin because it shared some similar physiochemical properties with the *T. brucei* EP procyclins, including surface

orientation, life cycle stage specificity, acidity, and immunodominance (Roditi *et al.,* 1989; Beecroft *et al.,* 1993; Bayne *et al.,* 1993; Garside *et al.,* 1995).

```
1   ctcgcactcctactccaagccagcaagaagcgtgaataccatgacgacaaccatgtcccg
                                                M  T  T  T  M  S  R
61  tgtcctacacctaatgacggtcacactgctgtgcgcacgcgtgggaatgggccaggctag
     V  L  H  L  M  T  V  T  L  L  C  A  R  V  G  M  G  Q  A  S
121 cgatgacgacgattgtggcgggcagagcattccccaaaaggtggaggaggtgcaaacgat
     D  D  D  D  C  G  G  Q  S  I  P  Q  K  V  E  E  V  Q  T  M
181 gtgcgacgttgcgcggcagctgagggccctggagaccgcttcccagtccgcggtggctgc
     C  D  V  A  R  Q  L  R  A  L  E  T  A  S  Q  S  A  V  A  A
241 cgtggtttcttctgcccgggaggcgtcggaggcgaaagagcgtgcggagaaagctgtgga
     V  V  S  S  A  R  E  A  S  E  A  K  E  R  A  E  K  A  V  E
301 gcgcgccaaatcgaagaagcgtggtgtggacgcggcgacggaagcggctgcaagggctgc
     R  A  K  S  K  K  R  G  V  D  A  A  T  E  A  A  A  R  A  A
361 ggctgcggcccagcgcgcggagacggtggtgagcgatgcgaggaagcacgcggcagacct
     A  A  A  Q  R  A  E  T  V  V  S  D  A  R  K  H  A  A  D  L
421 gacggcggcgtcgaaggatgctatcgagacgaccgacgagtcgctgcgcctactggccac
     T  A  A  S  K  D  A  I  E  T  T  D  E  S  L  R  L  L  A  T
481 atgcgagaaagcggacgagcccatccgcactgctgcaaaaaagtgcacgggtgccgccgc
     C  E  K  A  D  E  P  I  R  T  A  A  K  K  C  T  G  A  A  A
541 cgaagtcacgtccaagtcccttgagtcagcgttcgacgctctcgcggaactgctaccgga
     E  V  T  S  K  S  L  E  S  A  F  D  A  L  A  E  L  L  P  D
601 tggtgcggacgacatccgcgagcacggtgccgtgttcgtgaaggggctgaagtctctgga
     G  A  D  D  I  R  E  H  G  A  V  F  V  K  G  L  K  S  L  E
661 ggatgacgtgcgcacggctggagaggcaaagagcgaggcggagaaggctgagggcgatgc
     D  D  V  R  T  A  G  E  A  K  S  E  A  E  K  A  E  G  D  A
721 gaacgacgcggcagatggtgcccgtgccgtgctgacgggcgtgtgcgtgctgctgcttct
     N  D  A  A  D  G  A  R  A  V  L  T  G  V  C  V  L  L  L  L
781 ggctgcactgcacttttctgcggggctgtgaggcgggatcccccagctcacctgcccgcg
     A  A  L  H  F  S  A  G  L
841 agacattatg cgtattttga gtgacgaact tttattcgcc caacctggca taatacctta
```

Figure 5.1. The gene and translated protein sequence of glutamic acid/alanine-rich protein (GARP).

The DNA sequence of GARP (accession number L08055). The numbered lines contain the nucleotide sequence and immediately beneath the nucleotide sequence is the deduced amino acid sequence. In the protein sequence a potential N-terminal and C-terminal hydrophobic leader peptide and hydrophobic tail are indicated in bold (modified with permission from Bayne, et al. 1994).

GARP is a 256 amino acid polypeptide with a potential N-terminal leader sequence and a C-terminal hydrophobic sequence that allows GPI anchor attachment (Bayne *et al.,* 1993). There are no potential *N*-linked glycosylation sites (N-*X*-(S/T)), but the molecule appears to be heavily glycosylated as indicated by the much higher apparent molecular

mass than predicted based on the amino acid sequence: 40-44 kDa instead of 21.6 kDa, suggesting a high level of post-translational modification. The 21.6 kDa polypeptide component carries two large mannose- and galactose-containing oligosaccharides associated, presumably, by phosphodiester linkage, with threonine residues (Thomson *et al.,* 2002). The GARP GPI anchor contains an *sn*-1-stearoyl-2-L-3-HPO4-1-(2-*O*-acyl)-D-*myo*-inositol phospholipid moiety, that is resistant to the enzyme PI-phospholipase C due to a fatty acid found esterified to the 2-OH of the inositol ring (Thomson *et al.,* 2002) (Figure 5.2). GARP is a relatively abundant surface molecule with a minimum copy number of 2×10^5 copies/cell (Thomson *et al.,* 2002), which is significantly lower than the 2×10^6 copies of EP procyclin/cell on *T. brucei* procyclic culture forms.

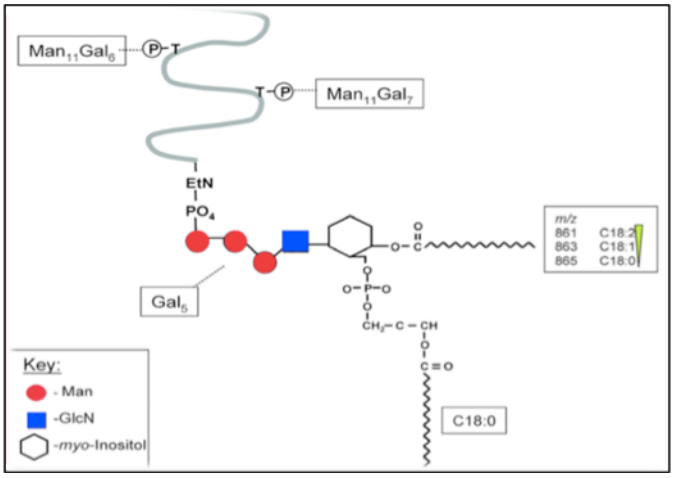

Figure 5.2. Model of GARP structure.
Schematic representation of GARP showing the PI moiety of the GPI anchor, proposed GPI anchor, and phosphosaccharide side chains (modified from Thomson, et al., 2002; with Permission: M. Ferguson).

When *T. brucei* BSF differentiate to PF in culture or in the tsetse, EP and GPEET surface procyclins are co-induced (Bütikofer *et al.,* 1999; Vassella *et al.,* 2000). Once induced, EP expression is maintained, whereas GPEET is repressed after seven to nine days

in the fly (Vassella *et al.,* 2000). Unlike the *T. brucei* EP procyclin, GARP was found to be developmentally regulated, as with GPEET procyclin, rather than being constitutively expressed (Bütikofer *et al.,* 2002). In *T. congolense*, GARP is briefly expressed shortly after BSF differentiation to PF and, much like procyclin, is present during the displacement of VSG (Roditi *et al.,* 1989; Beecroft *et al.,* 1993). GARP, largely absent from established procyclic forms in the fly midgut, is upregulated in epimastigotes (Bütikofer *et al.,* 2002). Since GARP is down-regulated in established procyclic midgut forms it is no longer called the *T. congolense* procyclin. A subsequently discovered molecule, TcHRP, is now deemed to be the *T. congolense* procyclin (Utz *et a.*, 2006). The Bütikofer lab in Bern, Switzerland infected flies with procyclic forms of both kilifi- and savanna-type *T. congolense* strains to study protein expression. It is possible that other clones of *T. congolense* show different expression profiles and it is also worth mentioning and emphasizing that the flies were infected with procyclic forms, not BSF as would be the case *in vivo.* This may be the reason that they only detected very low levels of GARP on the surface of early procyclic forms. In PCF used in the Pearson lab at UVic, GARP is expressed at an intermediate level and different expression levels are seen in different clones.

GARP is interesting because of its differential expression, appearing early in the transformation of BSF to PF, during and after VSG displacement. This coexistence of VSG and GARP on the parasite surface implies some very interesting functional and structural aspects to GARP that have been proposed but never studied at a structural level. In the case of *T. brucei* spp. it was shown that BSF are resistant to lysis by host serum components whereas PF or BSF whose VSG coat has been digested with trypsin are susceptible to complement-mediated lysis by the alternative pathway (Vickerman, 1985). If these active

complement factors, taken up in the bloodmeal, were to remain active for any length of time in the tsetse midgut, they could have deleterious consequences for trypanosomes that shed the VSG coat, in exchange for procyclin, too rapidly. The simultaneous expression of both VSG and procyclin on the surface would therefore confer a selective advantage during differentiation. While VSG protects the trypanosome from active complement in the bloodmeal, procyclins provide a highly acidic protease-resistant coat that is proposed to help the parasite survive in the harsh environment of the tsetse midgut (Thomson *et al.,* 2002).

Even though GARP is no longer considered the *T. congolense* procyclin, it seems to share many important features with procyclin. Like procyclin, GARP has been predicted to form an extended rodlike structure with potential for alpha helix formation (Roditi *et al.,* 1989; Beecroft *et al.,* 1993; Bayne *et al.,* 1993), is highly resistant to several proteinases (Bayne *et al.,* 1993, unpublished data), and appears during the transformation from BSF to PF (Beecroft *et al.,* 1993), in addition to surface orientation, immunodominance and acidity. Therefore the question becomes: is GARP, like procyclin, a protein that is expressed by trypanosomes upon entering the fly in order to mitigate the loss of VSG and ease the transition between the two disparate environments? To shed light on this question, I derived and characterized two GARP-specific monoclonal antibodies and used them for monitoring GARP expression during the *T. congolense* life cycle. In addition I determined the sequence and location of their polypeptide epitopes using mass spectrometry and crystallography, respectively. These epitopes were assigned to the crystal structure of GARP (in collaboration with Dr. Martin Boulanger's lab; UVic Biochemistry Dept) allowing us to determine the probable orientation of GARP on the trypanosome membrane.

Chapter 5.2: Materials and Methods

Two BALB/c mice were immunized with recombinant GARP in preparation for monoclonal antibody generation. Unfortunately one mouse died after administration of the final boost, so only one mouse was available for the fusion. Fusion was performed using splenocytes from the single spleen and monoclonal antibodies derived and screened, and analyzed according to outlined protocols in Chapter 2.

5.2.1. Cloning, expression and purification of rGARP

(work performed in Dr. M. Boulanger's laboratory by Jeremy Mason)

Forward and reverse primers were designed for amplification of a truncated protein construct that lacked the hydrophobic N-terminal signal peptide and C-terminal GPI-anchor sequence. BamHI and EcoRI sites were included in the forward and reverse primers, respectively, and PCR products were sub-cloned into the pGEX4T-1 plasmid (GE Lifesciences) which incorporates an N-terminal glutathione S-transferase (GST)-tag and thrombin cleavage site. Since *T. congolense* possesses several genomic copies of the GARP gene sequence, which are devoid of introns, genomic DNA could be used for PCR-amplification. Sequence analysis of numerous sub-cloned plasmids revealed a small number of polymorphisms within the gene and a prominently isolated clone was selected for protein expression.

Recombinant GARP was produced in *E. coli* BL21 cells and purified using glutathione Sepharose 4B beads (GE Lifesciences). The protein was eluted from the beads using 10 mM reduced glutathione and the GST-tag was subsequently removed by thrombin cleavage.

Further purification was performed by size exclusion and ion-exchange chromatography with fractions analyzed by SDS-PAGE and pooled based on purity.

5.2.2. Crystallization and data collection

(M. Boulanger's laboratory protocol)

Large diffraction quality crystals were grown overnight in 12.5 % PEG-1500 using the sitting drop vapor diffusion method. Drops consisted of 0.6 µl protein (35 mg/ml) with 0.6 µl reservoir solution and were equilibrated against 100 µl of reservoir solution. Cryo-protection was carried out in mother liquor supplemented with 12.5 % glycerol and 0.5 M sodium iodide for 50 seconds and flash cooled to 100 K directly in the cryo stream. Diffraction data were collected on a Rigaku R-AXIS IV++ area detector coupled to an MM-002 X-ray generator with Osmic 'blue' optics and an Oxford Cryostream 700. Diffraction data to 1.5 A° were processed using CrystalClear/ d*TREK (Pflugrath, 1999).

5.2.3 Epitope mapping of selected monoclonal antibodies

GARP was prepared for in-solution digestion with either trypsin or Glu-C by precipitating 50 µg per reaction of rGARP overnight at 4 °C by using a 9:1 volume ratio of ice-cold acetone to protein. Precipitates were pelleted by centrifugation at 10,000 rpm (7378 x g) for 1 min on a SIGMA 1-14 microcentrifuge (Osterode, Germany). Acetone was removed and the protein resolubilized in 30 µL of the appropriate buffer: 30 µL 25 mM ammonium bicarbonate in 6 M urea (for trypsin digestion), or 30 µL 25 mM ammonium carbonate pH 7.8 (for Glu-C digestion). For analysis by mass spectrometry, samples were reduced with 30 µL 10 mM dithiothreitol (DTT) for 45 min at 37 °C, alkylated with 30 µL

40mM Iodoacetamide for 45 min at 37 °C in the dark and the reactions quenched by addition of 30 µL 40 mM DTT. Following reduction and alkylation, rGARP was digested, using an enzyme to protein ratio of 1:20, overnight at 37 °C with trypsin (20 ng/µL stock, #V5111, Sequencing grade modified porcine trypsin, Promega, Madison, WI, USA), or at room temperature with Glu-C endoproteinase (1 µg/µL stock, #90054, MS grade, Thermo Scientific, Rockford IL, USA). To stop digestion 100x protease inhibitor cocktail set V, EDTA-free (#539137, Calbiochem®, La Jolla, CA, USA) was added to a final concentration of 1x. All digests were diluted with PBS to give a final concentration of 0.2 µg GARP/µL.

Immunocapture mass spectrometry was used to identify the peptide epitopes bound by monoclonal antibodies. The murine monoclonal antibodies were first captured on magnetic beads. These solid-phase adsorbents were used to capture peptide epitopes from the rGARP enzyme digests. Beads chosen for capturing the monoclonal antibodies of interest were goat anti-mouse IgG dynabeads (#110.33, Invitrogen Dynal AS, Oslo, Norway). The desired volume of Dynabead slurry (25 µL of Dynabeads, $\sim 1 \times 10^7$ beads) was transferred to a microfuge tube and washed with three quick washes of 500 µL PBS each. Tubes were placed on a magnet (#2010, Dynal® Invitrogen bead separator) for 1 min and the supernatant discarded. According to the manufacturer's instructions, 25 µL of Dynabeads can capture 0.5 – 1.5 µg of primary antibody, therefore each 25 µL of beads were incubated with 1 mL of specific monoclonal antibody tissue culture supernatant (which contains ~ 1 – 10 µg/mL specific antibody) overnight at 4 °C on a rotator. An excess of mAb was used to ensure complete bead saturation. The following day tubes were placed on a magnet and mAb supernatant removed. The beads were washed three times with 1 mL PBS each. After the final wash, beads were resuspended in PBS to obtain the original volume, 25

µL, and stored at 4 °C. One aliquot of mAb 2-D7 beads was incubated with trypsinized GARP and another aliquot of mAb 2-D7 beads was incubated with Glu-C-digested GARP. The same was done for the mAb 4-B7 beads. Additionally, each digest was incubated with naked beads, beads without any primary antibody captured, in order to be able to assess for non-specifically bound peptides.

For each peptide capture reaction 50 µL of digested rGARP (0.2 µg/µL, ~10 µg) were added to 25 µL of mAb-coated Dynabeads. Capture reaction mixtures were incubated at room temperature for 1.5 hours with shaking. Following peptide capture, tubes were place on a magnet and supernatant removed and saved (unbound, non-epitope peptides). Beads were washed with three times with 1 mL PBS to remove unbound or weakly bound peptides. Fifty microlitres of 5 % acetic acid was added to the beads to release the bound peptides, and the mixture incubated for 5 minutes at room temperature. Finally, the supernatant was removed and saved for mass spectrometric identification of peptide epitopes.

5.2.4. MALDI-TOF mass spectrometry

Antibody-captured peptides and non-captured peptides from trypsin digestion, and from Glu-C digestion and complete trypsin and Glu-C digest as controls, were desalted and concentrated using ZipTip (C18 resin; P10, Millipore Corporation, Bedford, MA, USA). For each sample, 1 µL of the desalted, concentrated peptide mixture was mixed 1:1 with the matrix, alpha-cyano-4-hydroxycinnamic acid and spotted onto a Voyager, 100 position, stainless steel MALDI plate (Applied Biosystems, Foster City, CA, USA). Calmix 2 internal standards (purchased from UVic mass spec. facility), angiotensin 1 (1296.6853), ACTH 1-17

clip (2093.0867) and ACTH 18-39 clip (2465.1989) were used to calibrate the Voyager prior to analysis of samples. An Applied Biosystems Voyager DE-STR mass spectrometer (Applied Biosystems, Foster City, CA, USA) running in delayed extraction, reflectron mode was used to acquire MALDI-TOF data. The GARP protein sequences were submitted for *in silico* digestion using MS-digest (Protein Prospector software package; San Francisco, CA, USA: http://prospector.ucsf.edu/) and peptide masses were compared to the in-solution digests. In addition captured peptides were selected for tandem MS analysis to obtain sequence confirmation.

5.2.5. Tandem mass spectrometry (MS/MS)

Peptides from each sample of interest were spotted onto a 384 OptiT MALDI plate and analyzed on a 4800 MALDI-TOF/TOF (AB Sciex, Concord, ON, Canada) with explorer analyzer v3.5 installed to obtain sequence information.

5.3. Results

5.3.1. The selection and analysis of anti-GARP mAbs 2-D7 and 4-B7 by ELISA, immunoblot analysis, flow cytometry and immunofluorescence

Hybridoma supernatant screening by ELISA and immunoblotting on recombinant GARP resulted in the selection of 15 stable clones that were positive in both assays. To evaluate the ability of the mAbs to react with native GARP in the presence of post-translational modifications (PTMs) and to determine if they are species, sub-species and strain specific, I tested the reactive mAbs on immunoblots of proteins separated from

lysates of PCF from *T. congolense* IL 3000 (savannah strain), *T. congolense* k45/1 (Kilifi strain), *T. simiae* CP 11 and *T. b. brucei* 427. Of the 15 mAbs tested, five reacted with *T. congolense* IL3000 PCF only and one mAb, 2-D7, reacted with PCF of both *T. congolense* IL 3000 and *T. simiae* CP11. None of the mAbs reacted with proteins from *T. b. brucei* 427. The results of immunoblotting using two of the selected mAbs are shown in Figure 5.3. MAb 2-D7 was found to be subgenus *Nannomonas*-specific- it recognized antigens in *T. congolense* IL 3000, *T. congolense* K45/1, and *T. simiae* (Figure 5.3; Panel A), whereas mAb 4-B7 was found to be species and strain specific for *T. (Nannomonas) congolense* savannah only (Figure 5.3; Panel B).

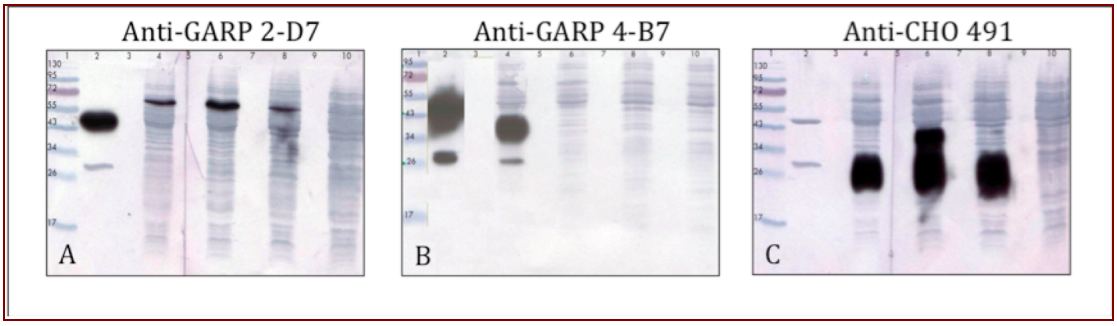

Figure 5.3. Immunoblot analysis of mAbs 2-D7 and 4-B7 on proteins from PCF lysates of T. congolense IL 3000.

Shown are the nigrosin-stained PVDF membranes with immunoblot film overlaid. The primary antibodies used were mAb 2-D7 (Panel A), mAb 4-B7 (Panel B) and mAb 491 (Panel C). Lane 1: pre-stained MW markers (kDa), lane 2: 0.4 µg of rGARP, lanes 4, 6, 8 and 10: lysates from 5×10^6 PCF of T. congolense IL 3000, T. congolense kilifi 45/1, T. simiae CP-11, and T. b. brucei 427 respectively.

What is interesting is that while mAb 2-D7 reacted with a band at approximately 45 kDa in the rGARP extract (GARP-GST fusion protein) it only recognizes a band at approximately 70 kDa in the lysates of *T. congolense* IL 3000 and K45/1 and *T. simiae* CP11 PCF (Figure 5.3 A; lanes 4, 6 and 8). On the other hand, mAb 4-B7 reacted with two bands at apparent masses of 43 and 26 kDa in the lysate of *T. congolense* IL 3000 and rGARP

extract (Figure 5.3 B; lanes 2 and 4). Most likely the band at 26 kDa in the *T. congolense* IL 3000 lysate is a GARP precursor or breakdown product, whereas the similar MW band in the rGARP extract is truncated GARP or GARP that has lost its GST tag. The protein band recognized by mAb 2-D7 has a mass (70 kDa) much higher than that reported for GARP in the literature (43 kDa). It is possible that this mAb recognizes a highly glycosylated form of GARP or a dimmer, or that this is a form of GARP that still has its GPI anchor attached. There is a small chance that this epitope may be specific to a particular *GARP* gene, since multiple *GARP* genes exist at two distinct loci (Rangarajan *et al.,* 2000). However, GARP is very highly conserved within the subgenus *Nannomonas* such that the polypeptide length is probably the same for all gene products. It is possible that, like procyclin, the negative charges on GARP and the presence of extensive glycosylation cause it to interact abnormally with SDS and to migrate aberrantly on gels (Richardson *et al.,* 1988). Included in the immunoblot analysis was the positive control mAb 491 (Figure 5.3; Panel C). This mAb fails to recognize the rGARP simply because in the *E. coli* expression system used, GARP is not glycosylated, confirming that mAb 491 recognizes a carbohydrate epitope (Mookherjee and Pearson, 2001).

Flow cytometry was used to determine the surface reactivity of all six mAbs on live PCF of *T. congolense* IL 3000, *T. simiae* CP11 and *T. b. brucei* 427. All mAbs, including 2-D7, were negative on PCF of both *T. simiae* CP 11 and *T. b. brucei* 427. Data indicate that mAb 2-D7 binds a linear epitope that is otherwise concealed under non-denaturing conditions, perhaps by a glycosylation or concealment within the tertiary structure of the GARP molecule expressed on *T. congolense* and *T. simiae*. Only one mAb, 4-B7, demonstrated reactivity with the surface of live PCF of *T. congolense* IL3000, indicating that although the

epitope is linear it is not concealed under non-denaturing conditions or within the tertiary structure-it is clearly a surface epitope. Based on binding patterns in ELISA, reactivity on immunoblots (Figure 5.3) and surface immunoreactivity on live cells (Figure 5.4), mAbs 2-D7 and 4-B7 were selected for more detailed study.

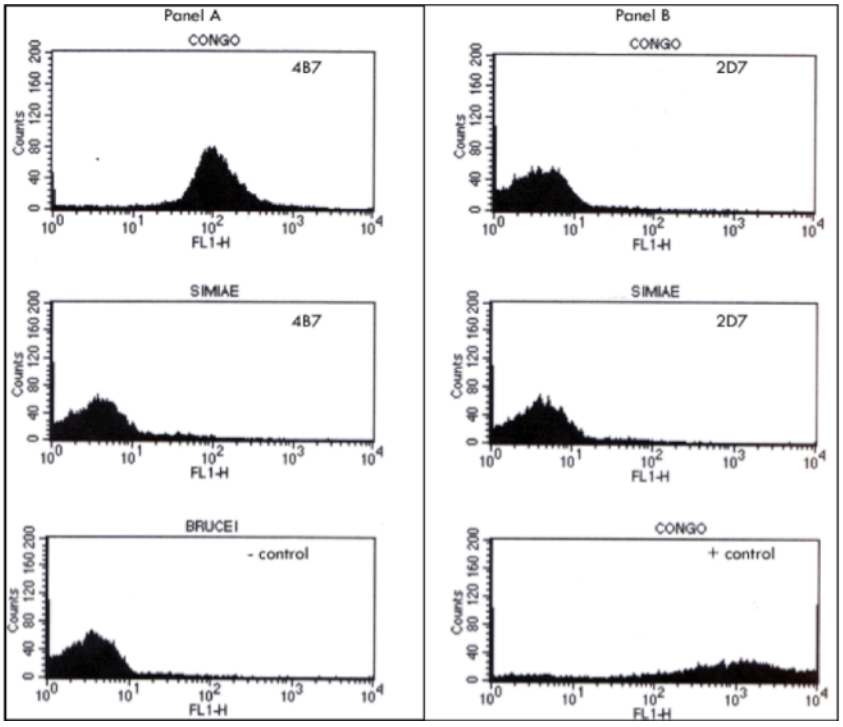

Figure 5.4. Flow cytometric analysis of anti-GARP mAbs 2-D7 and 4-B7 on live T. congolense IL 3000, T. simiae CP11 and T. b. brucei 427 PCF.

Anti-GARP mAb 4-B7 (Panel A) and mAb 2-D7 (Panel B) were tested for surface binding to T. congolense IL3000, T. simiae CP11 and T. b. brucei 427 PCF (negative control). Anti-(Nannomonas)-carbohydrate mAb 491 (Beecroft et al., 1993) was used on T. congolense IL3000 as a positive control. Undiluted tissue culture supernatants were used with a 1:50 dilution of fluorescein-conjugated goat anti-mouse IgG/IgM secondary antibody. For every sample, ~5×10^6 cells were analyzed.

Life cycle stage specificity was analyzed by immunoblotting on SDS-PAGE-separated proteins from the lysates of *T. congolense* IL 3000 PCF and EMF grown *in vitro* (Figure 5.5). Neither mAb was lifecycle stage specific; both react with *T. congolense* IL 3000 lysates. This indicated GARP to be present in both the PCF and EMF. Immunoblot analysis also showed

there to be more of GARP in EMF compared to the PCF, as seen by a thicker and darker band observed in lanes 3 and 5 of Figure 5.5. Once again mAb 2-D7 reacts with GARP at 70 kDa only, while mAb 4-B7 reacts with GARP at 26 and 40 kDa in EMF lysates.

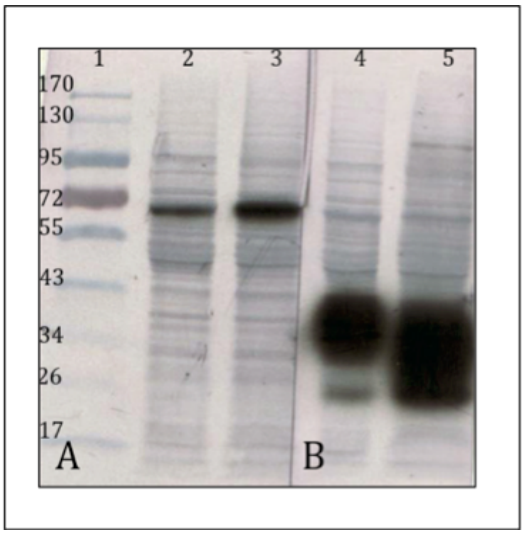

Figure 5.5. Immunoblot analysis of mAbs 2-D7 and 4-B7 on separated proteins from PCF and EMF lysates of T. congolense IL 3000 to determine life cycle stage specificity.

Shown are the nigrosin-stained PVDF membranes with immunoblot film overlaid. The primary antibodies used were mAb 2-D7 (Panel A), mAb 4-B7 (Panel B). Lane 1 contained pre-stained MW markers (kDa), lane 2 and 4, and lane 3 and 5 contained lysates from 2×10^6 PCF of T. congolense IL 3000, and T. congolense IL 3000 EMF, respectively.

Immunofluorescence microscopy was performed on live and acetone-fixed/permeabilized PCF and EMF of *T. congolense* IL 3000 to determine whether the epitope recognized by mAb 2-D7 would become exposed after treatment of the cells with acetone (Figure 5.6). Acetone dissolves the cell lipids and dehydrates the cells, while maintaining cellular and subcellular architecture, thus permitting unhindered access of antibodies to subcellular compartments. Anti-EP repeat mAb 247 (*T. brucei* spp. specific) was used as a negative control (Figure 5.6; Panel 3, A-D). This antibody binds only to EP

procyclin in *T. brucei* spp. The epitope recognized by mAb 4-B7 was not destroyed by acetone-fixation (Figure 5.6; Panel 1, A-D), albeit the fluorescence intensity was slightly weaker, especially in the fixed EMF, than that seen in the live cells. This mAb recognizes its epitope in both PCF and EMF, corroborating the immunoblot findings that this mAb is not life cycle stage-specific. Previous work has shown GARP to be upregulated in EMF, and in my work this increase can be seen by a darker staining protein band seen in the lysates from *T. congolense* IL 3000 EMFs (Figure 5.5; lanes 3 and 5), as well as a dramatic increase in surface fluorescence observed using mAb 4-B7 on the live EMF compared to PCF. Although mAb 2-D7 did not react to its epitope on live PCF, it was able to react weakly when the cells were acetone-fixed (Figure 5.6; Panel 2, A-D). This suggests that its linear epitope must be concealed by its position or arrangement of GARP in the cell membrane. Monoclonal antibody 2-D7 was also reactive with both life cycle stages. Anti-*T. b. brucei* mAb 247, used as a negative control, did not show fluorescence with acetone-fixed or live *T. congolense* PCF and EMF (Figure 5.6; Panel 3, A-D).

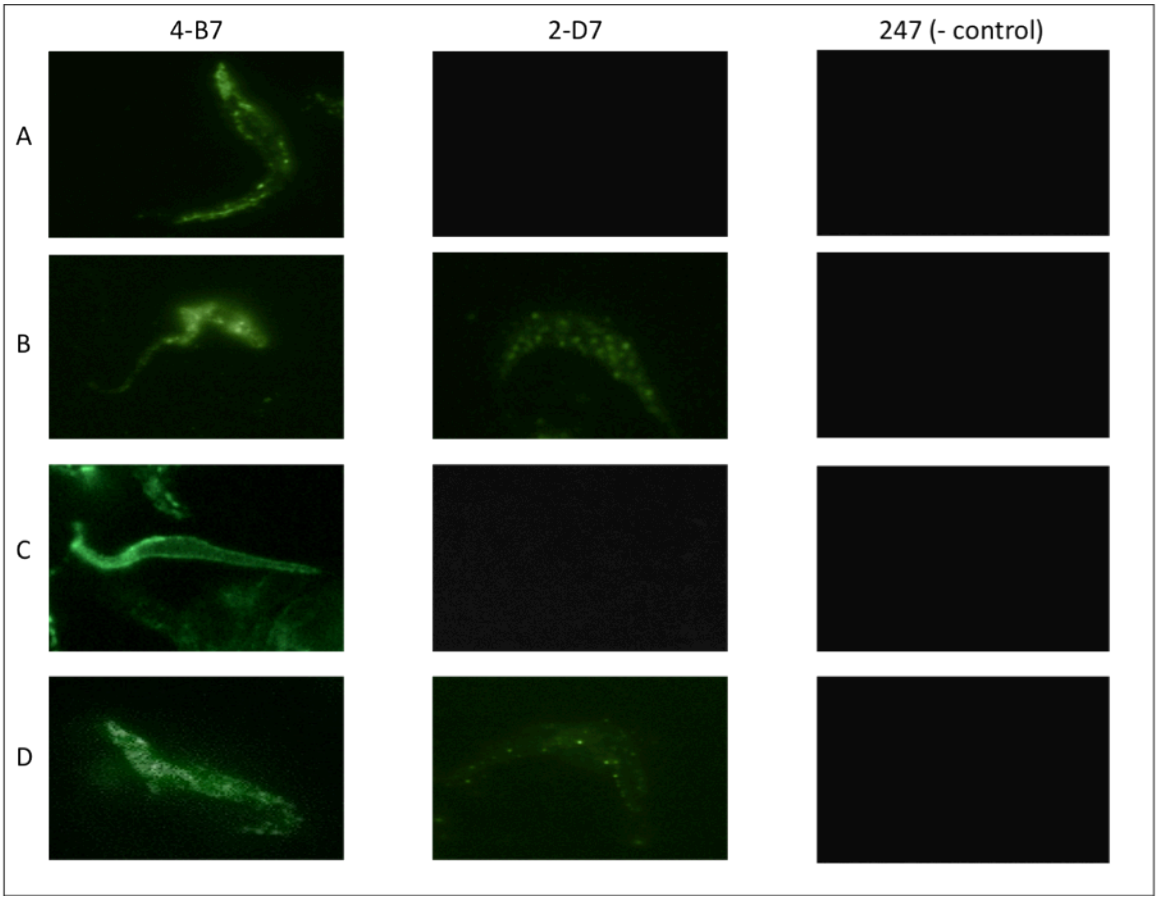

Figure 5.6. Immunofluorescence analysis of mAbs 4-B7 and 2-D7 on live and acetone-fixed T. congolense PCF and EMF grown in vitro.

Horizontally - A: Live PCF. B: acetone-fixed PCF. C: Live EMF. D: acetone-fixed EMF. Vertically - Panel 1: Anti-GARP mAb 4-B7. Panel 2: Anti-GARP mAb 2-D7. Panel 3: Anti-EP repeat of T. b. brucei EP procyclin. Undiluted tissue culture supernatants containing primary antibodies were used, followed by a 1:50 dilution of Alexafluor 488 conjugated goat anti-mouse IgG/M secondary antibody.

Life cycle stage morphology of the procyclic and epimastigote culture forms used in the above experiments was verified using the DNA-binding dye DAPI to localize the kinetoplast and nucleus (Figure 5.7). Typically, the nucleus in PCF is located centrally, with the kinetoplast posterior to the nucleus (Figure 5.7; panel A). As trypanosomes differentiate into epimastigote forms they become longer and the kinetoplast migrates anterior to the nucleus, towards the flagellum (Figure 5.7; panel B). This elongation and

anterior migration of the kinetoplast is accompanied by biochemical and structural changes that prepare the parasite for differentiation into the mammal-infective MCF. My DAPI analysis of the procyclic and epimastigote culture forms is consistent with the published literature showing that the forms I grew in culture were truly PCF and EMF.

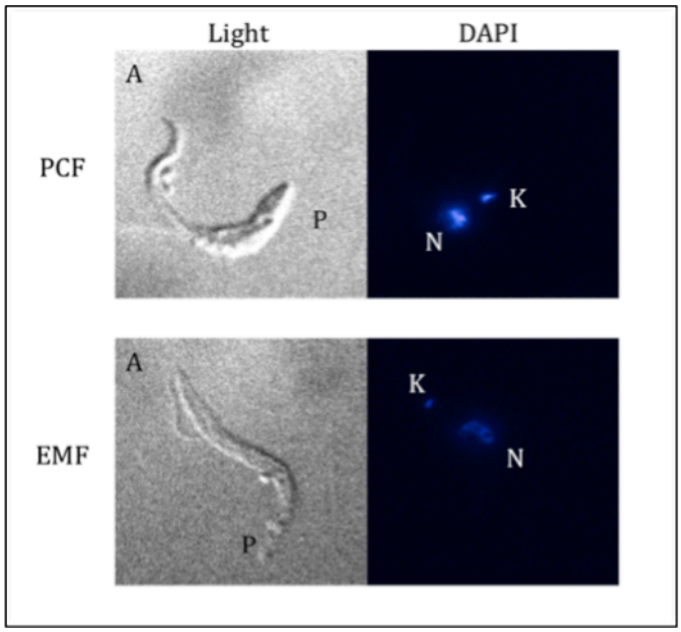

Figure 5.7. DAPI counterstained images of acetone-fixed T. congolense PCF and EMF grown in vitro.

A: PCF B: EMF. The kinetoplast (K) and nucleus (N) are indicted, and the posterior (P) and anterior (A) ends of the trypanosomes have been labeled. It is counterintuitive that the flagellum points to the anterior end of the trypanosomes.

To continue the characterization of GARP, and to validate the results obtained on parasites grown *in vitro*, I tested mAb 4-B7 on *T. congolense* parasites taken directly from different compartments of infected tsetse. Dr. Lee Haines at the Liverpool School of Tropical Medicine, Liverpool, UK, supplied slides of air-dried parasite samples (Figure 5.8).

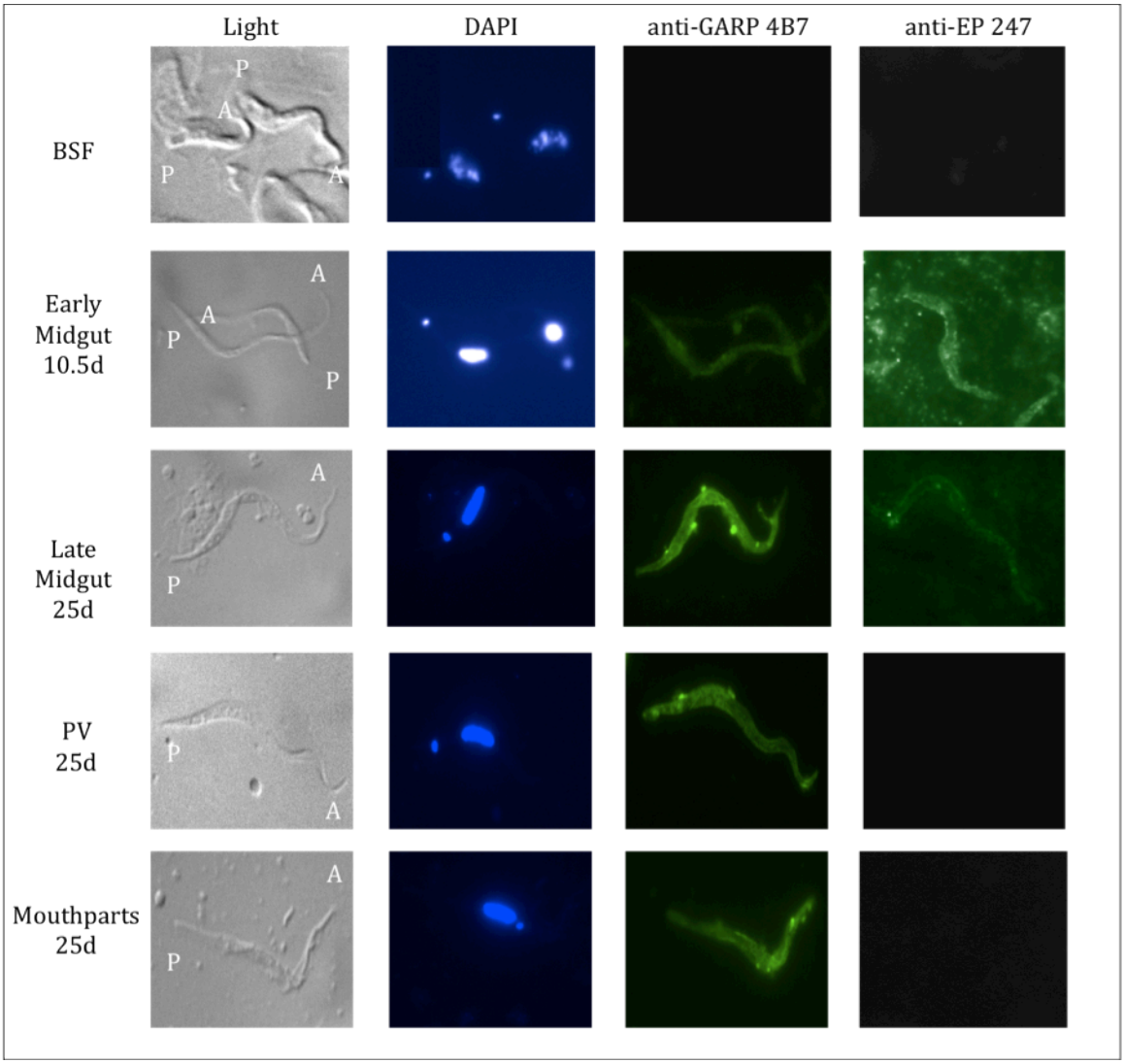

Figure 5.8. Immunofluorescence analysis of GARP expression on T. congolense BSF and insect forms taken directly from infected rats and tsetse, respectively.
T. congolense 1/148 savannah BSF isolated from infected rats (first horizontal panels), procyclic forms isolated from the midguts of infected tsetse after 10.5 and 25 days (second and third horizontal panels), proventricular (PV) forms (fourth horizontal panels) and epimastigote forms isolated from the mouthparts (fifth horizontal panel). Trypanosomes were simultaneously labeled with anti-GARP 4-B7 and the DNA-binding dye DAPI. The posterior (P) and anterior (A) ends of the trypanosomes are indicated. A negative control antibody, anti-T. b. brucei EP repeat mAb 247, was included with all life cycle stages.

Immunofluorescence results for trypanosomes taken from infected tsetse are consistent with those obtained from trypanosomes grown *in vitro.* This mAb reacted weakly to early midgut forms, which are reported to express low levels of GARP. Strong

fluorescence was seen with the late-stage established midgut forms, which are probably ready for transit to the mouthparts and subsequent differentiation into adherent EMF. The same negative control antibody, anti-EP procyclin mAb 247, was used. This mAb shows very high background in the midgut, especially with the early midgut forms (Figure 5.8; 2nd and 3rd horizontal panel). It has been recently observed that tsetse, specifically *Glossina morsitans morsitans* (the species used in this work) and *G. palpalis palpalis,* express a tsetse-specific midgut molecule, tsetse-EP, that acts as a powerful antagonist of trypanosome establishment in the midgut and is thought to interact with the trypanosomes themselves (Haines *et al.,* 2010). The EP repeat section of this tsetse molecule is similar to the repeat section of the *T. b. brucei* procyclin EP (Vassella *et al.,* 2001), explaining the cross-reactivity of mAb 247 with *T. congolense*. The tsetse EP protein is an immunoresponsive molecule that is upregulated when tsetse are immunogen-challenged, thus it would be expected that the early midgut preparations would have more of this tsetse-EP compared to the late midgut preparations and as a result show greater reactivity with mAb 247. Neither the PV nor the mouthpart preparations showed any cross-reactivity with mAb 247, consistent with this protein either not being produced in the PV or mouthparts, being rapidly turned over or rapidly translocated to the midgut (Haines *et al.,* 2010; Chandra *et al.,* 2004).

A summary of the data obtained for mAbs 2-D7 and 4-B7 is presented in Table 5.1.

Table 5.1. Summarized data: characterization of selected anti-GARP monoclonal antibodies.

	mAb 2-D7	**mAb 4-B7**
Isotype	IgG_2a	IgG_2a
Species/ strain[1] specificity	*Nannomonas*-specific[2]	*T. congolense* specific[3]
Lifecycle stage[4]	PCF and EMF	PCF and EMF
Flow cytometry[5]	Internal	Surface
Immunofluorescence microscopy[6] EMF	 -	 +4
Immunofluorescence microscopy[7] PCF (grown *in vitro*) EMF (grown *in vitro)*	 +1 +1	 +3 +3
Immunofluorescence microscopy[8] BSF (from rat blood) Early midgut (from tsetse) Late midgut (from tsetse PV (from tsetse) Mouthparts (from tsetse)	 Not done	 - +1 +4 +3 +3

[1]Determined by immunoblot analysis on parasite lysates.
[2]Reacts with proteins in lysates of *T. simiae* CP11 as well as *T. congolense* IL3000 (Savannah) and K45/1 (Kilifi).
[3]Reacts with proteins in lysates of *T. congolense* IL 3000 and 1/148 strains, not to *T. simiae*).
[4]Determined by immunoblot analysis on *T. congolense* IL 3000 grown *in vitro.*
[5]Trypanosomes were intact PCF of *T. congolense* IL3000, *T. simiae* CP11 or *T. b. brucei* 427. Positive mAb binding was labeled as surface. Neither mAb reacted with *T. b. brucei*.
[6]Performed on intact, live EMF self-adhered to slide covers. Fluorescence intensity was graded as very strong (+4) to weak (+1) or negative (-).
[7]Epifluorescence microscopy on acetone-fixed *T. congolense* IL 3000 PCF and EMF grown *in vitro.*
[8]Epifluorescence microscopy on *T. congolense* 1/148 trypanosomes (*from tsetse*) bound to poly-L-lysine coated slides. Only mAb 4-B7 was tested.

5.3.2. Identification and localization of anti-GARP 2-D7 and 4-B7 mAb epitopes by in-solution digestion followed by MALDI-TOF and MALDI-TOF-TOF mass spectrometry.

With my pool of anti-GARP mAbs narrowed down to two: mAb 2-D7 (specific for a non-surface disposed epitope) and mAb 4-B7 (specific for a surface-exposed epitope), I next wanted to identify the amino acid sequences of the relevant epitopes, and localize their positions on GARP in an attempt to determine the orientation of the GARP molecule in terms of its disposition on the trypanosome surface membrane. To do this required a GARP crystal structure. We were fortunate to establish collaborations with Dr. Noboru Inoue and Tatsuya Sakurai (Hokkaido, Japan), and with Dr. Martin Boulanger and Jeremy Mason (Department of Biochemistry and Microbiology, UVic). Drs. Inoue and Sakurai provided a plasmid construct containing part of the GARP sequence for recombinant expression and Dr. Boulanger and Jeremy Mason performed the expression, purification and crystallization of GARP. Jeremy expressed and purified the rGARP-GST fusion protein and removed the GST tag, according to the Boulanger lab protocol. This material was used for crystallization and structure determination, and for my epitope localization experiments. The work to be described below has been written up for publication in the Journal of Biological Chemistry (Loveless et al., 2010).

Immunoblotting experiments were performed to evaluate the reactivity of mAb 2-D7 and 4-B7 to the rGARP in order to verify its successful expression, purification and cleavage. Both mAbs reacted very strongly with the rGARP, and showed a single band at approximately 26 kDa (Figure 5.9; lanes 5 and 10) rather than 26 and 45 kDa (Figure 5.3 A, B and C; lane 2) as had originally been seen using uncleaved rGARP fusion protein produced in Japan. The MW of the new rGARP produced in the Boulanger lab better reflects the

predicted molecular mass of the GARP polypeptide. The original uncleaved rGARP extract contained the 45 kDa rGARP-GST fusion protein, and the 26 kDa mass is likely a breakdown product of this fusion protein. Since immunizations had been performed using the original rGARP-GST fusion protein obtained from Japan, I included a lane with GST alone to investigate whether mAb 2-D7 and 4-B7 recognized the GARP polypeptide or the GST portion of the fusion protein. The results are shown in Figure 5.9. Fortunately, neither of these two anti-GARP mAbs bound to GST and bound to the GARP polypeptide as expected (Figure 5.9; lanes 2, 5, 7 and 10).

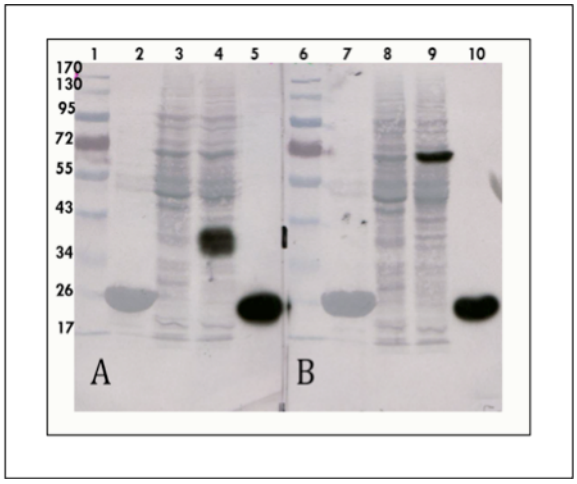

Figure 5.9. Immunoblot analysis of mAbs 2-D7 and 4-B7 on rGARP expressed, purified and cleaved in the Boulanger lab.

Shown are the nigrosin-stained PVDF membranes with the immunoblot film overlaid. Panel A: 4-B7. Panel B: 2-D7. Lanes 1 and 6 contained protein pre-stained MW markers (kDa), lanes 2 and 7 contained 1/8 dilution of GST stock, lanes 3 and 8 contained PCF lysate from 5×10^6 T. b. brucei 427 cells, lanes 4 and 9 contained PCF lysate from 5×10^6 T. congolense IL 3000 and lanes 5 and 10 contained 3 µg of rGARP.

Based on the above results I was confident that GARP had been successfully expressed, purified and cleaved and thus I continued with mAb 2-D7 and 4-B7 epitope determination. One can perform epitope mapping by two relatively simple methods that

have been developed for rapid and sensitive identification of epitope-containing peptides, based on direct MALDI-MS/MS analysis of epitope-containing peptides bound to affinity beads. In the first method the protein is cleaved into peptides prior to immunoenrichment of epitope-containing peptides (Raska *et al.,* 2003). This method requires that the epitope-containing peptide not be destroyed by cleavage of the protein. The second method, referred to as "protein footprinting" (Sheshberadaran and Payne, 1988), is analogous to DNase footprinting whereby the antigen is bound by antibody forming antigen-mAb complexes prior to proteolytic cleavage. The antigen-binding domains of an antibody are resistant to proteolysis (Porter, 1959; Parham, 1983), thereby conferring protection against degradation of the epitope region of the bound antigen (Jemmerson and Paterson, 1986; 1986). For mAbs 2-D7 and 4-B7epitope mapping I utilized the first method. GARP was cleaved with either trypsin or Glu-C and peptides captured using goat anti-mouse IgG Dynabeads with either mAb 2-D7 or 4-B7 capture antibodies bound via their Fc portions. These beads have two benefits: first, using beads coated with goat anti-mouse IgG antibodies allowed the selective capture of my antibodies from tissue culture supernatant that contained a high concentration of fetal bovine serum (and thus bovine immunoglobulins) and second, the magnetic Dynabeads ensured rapid concentration and minimal loss of sample. Non-specific binding of peptides to the affinity beads was analyzed using naked goat anti-mouse IgG Dynabeads. Beads with no capture antibody are notoriously sticky so these represented a very conservative background control. Figures 5.10 and 5.11 show the MALDI-TOF spectra for mAb 2-D7 and mAb 4-B7, respectively.

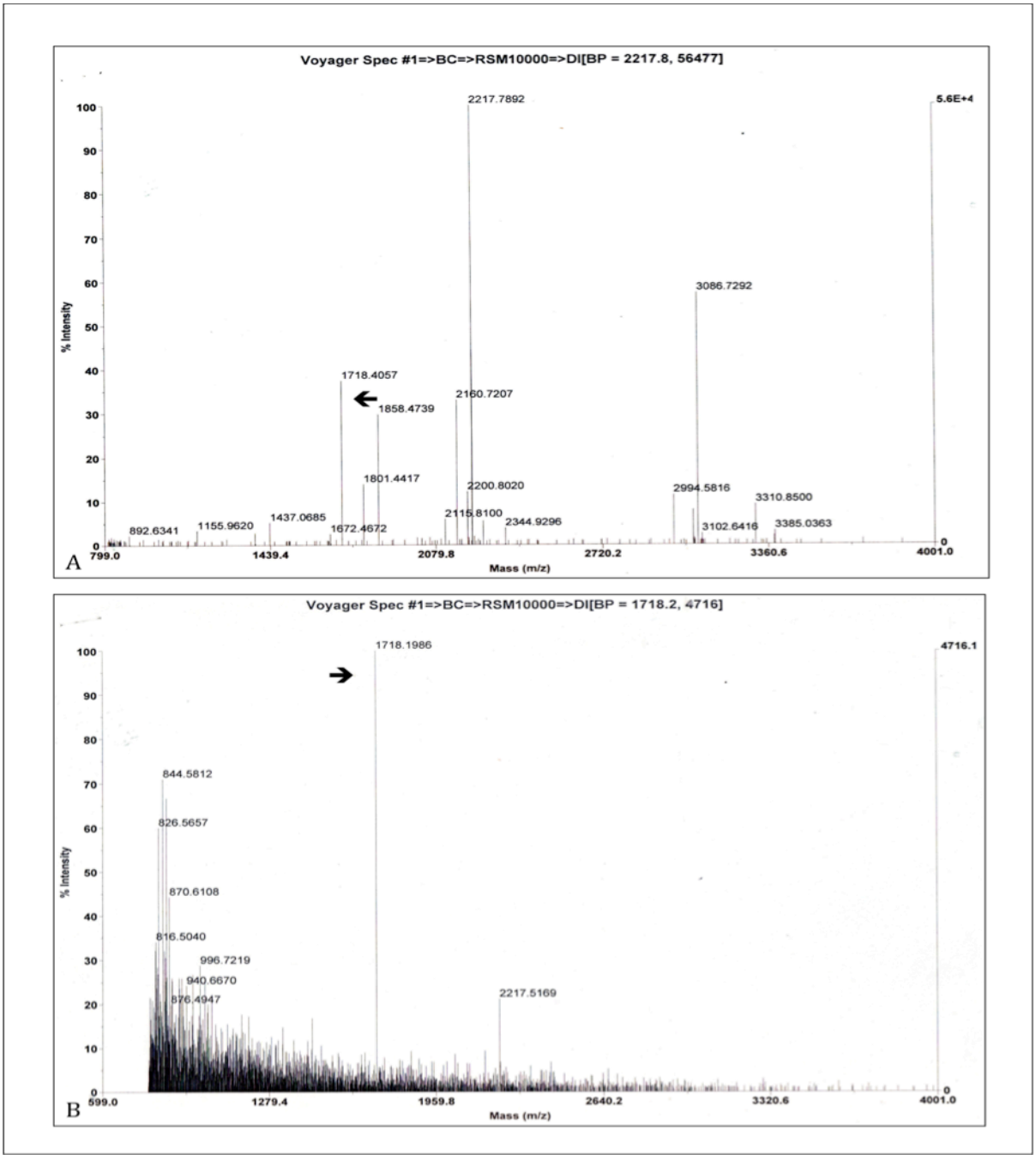

Figure 5.10. MALDI-TOF mass spectra of peptides in trypsin-digested GARP and peptides bound by mAb 2-D7.

Panel A: GARP trypsin digest (control). Panel B: mAb 2-D7 epitope peptide enrichment. GARP was digested with the serine protease trypsin to generate peptides that were then incubated with magnetic Dynabeads coated with mAb 2-D7. Panel A shows the spectrum of GARP tryptic peptides indicating a successful digestion. Panel B shows the enriched peptide mass: 1718 [M+H⁺] (indicated with an arrow).

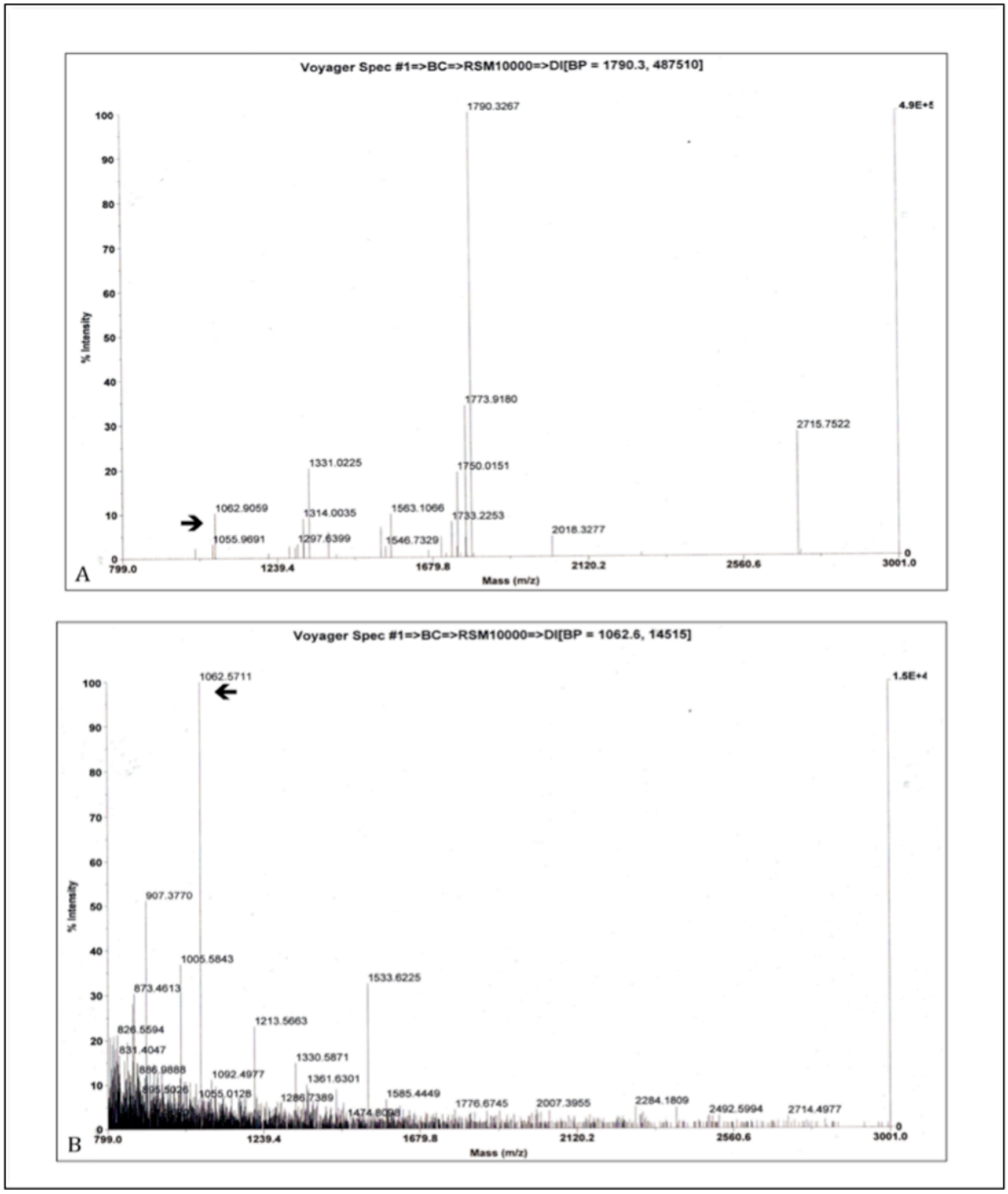

Figure 5.11. MALDI-TOF mass spectra of peptides in Glu-C digested GARP and peptides bound by anti-GARP mAb 4-B7.

Panel A: GARP GluC digest (control). Panel B: mAb 4-B7 epitope peptide enrichment. GARP was digested with the serine protease GluC to generate peptides that were then incubated with magnetic Dynabeads coated with mAb 4-B7. Panel A shows the spectrum of GARP GluC peptides indicating a successful digestion and Panel B shows the enriched peptide mass: 1062 [M+H⁺] (indicated with an arrow).

Epitope mapping experiments were performed on two separate occasions and both times the results obtained were very positive. Both the trypsin and GluC digestion of GARP was good (Figure 5.10 and 5.11; Panel A), there was an excellent array of masses that corresponded well to predicted peptides determined by the *in silico* digests of GARP. Mass fingerprint analysis of trypsin-digested GARP incubated with mAb 2-D7 beads identified a tryptic peptide having a monoisotopic mass [$M+H^+$] of 1718 that was selectively enriched (Figure 5.10; Panel B). Based on the *in silico* trypsin digest the corresponding amino acid sequence of the peptide of that mass was determined to be ^{45}ALETASQSAVAAVVSSAR^{62}. Thomson *et al.* previously observed this same peptide in their analysis of GARP tryptic digests by MALDI-TOF in which they determined that more than one *GARP* gene was expressed simultaneously (Thomson *et al.,* 2002). Compared to the trypsin digest spectrum (Figure 5.10; Panel A), the epitope enrichment worked extremely well and was very clean. The 1718 Da peptide was the major peptide at 4.7×10^3 peak intensity, with a minor peptide at 2217 Da, present at 9.43×10^2 peak intensity (Figure 5.10; panel B). Their peak intensities in the trypsin digest were 2.22×10^4 and 5.6×10^4 peak intensity, respectively (Figure 5.10; Panel A). The 2217 Da peptide was the most abundant peptide in the digest sample, resulting in signal saturation. The peak intensity of the 1718 Da peptide was lower in the enrichment sample (4.7×10^3) than in the digest sample (2.22×10^4). This may be the shortfall of using hybridoma tissue culture supernatants, which are known to have low concentrations of specific antibody, compared to ascites or purified antibody. I would, therefore, be unable to pull out all of this peptide from the digest sample. I was chiefly concerned with enrichment rather than signal saturation. No peptide enrichment occurred

when this same digest was incubated with mAb 4-B7. Presumably the epitope was destroyed by cleavage with trypsin (spectrum not shown).

Conversely, when GARP was digested with Glu-C and incubated with mAb 4-B7, a peptide having a monoisotopic mass [$M+H^+$] of 1062 Da was selectively enriched (Figure 5.11; Panel B). The 1062 Da peptide was the major peptide at 1.5×10^4 peak intensity, with a minor peptide at 1533 Da at 4.8×10^3 peak intensity (Figure 5.11; panel B). Once again, being principally concerned with enrichment, I was most interested in obtaining a clean spectrum with respect to the digest (Figure 5.11; Panel A); the 1062 Da peptide in the digest sample had a peak intensity of 4.9×10^4. The same digest when incubated with mAb 2-D7 beads resulted in no peptide enrichment: its epitope was destroyed by Glu-C digestion (spectrum not shown). Based on the *in silico* GluC digest, the corresponding amino acid sequence for the 1062 Da peptide was determined to be: ^{141}SLRLLATCE^{149}.

To be sure that neither of these two peptides, 1062 Da and 1718 Da, had bound non-specifically to the beads rather than being enriched by their respective antibodies, I incubated the digest samples with naked beads - beads without any capture mAb. Figure 5.12 shows the spectrum for the tryptic-GARP peptides that bound non-specifically (Figure 12; Panel A) and the spectrum for the GluC-GARP peptides that bound non-specifically (Figure 12; Panel B).

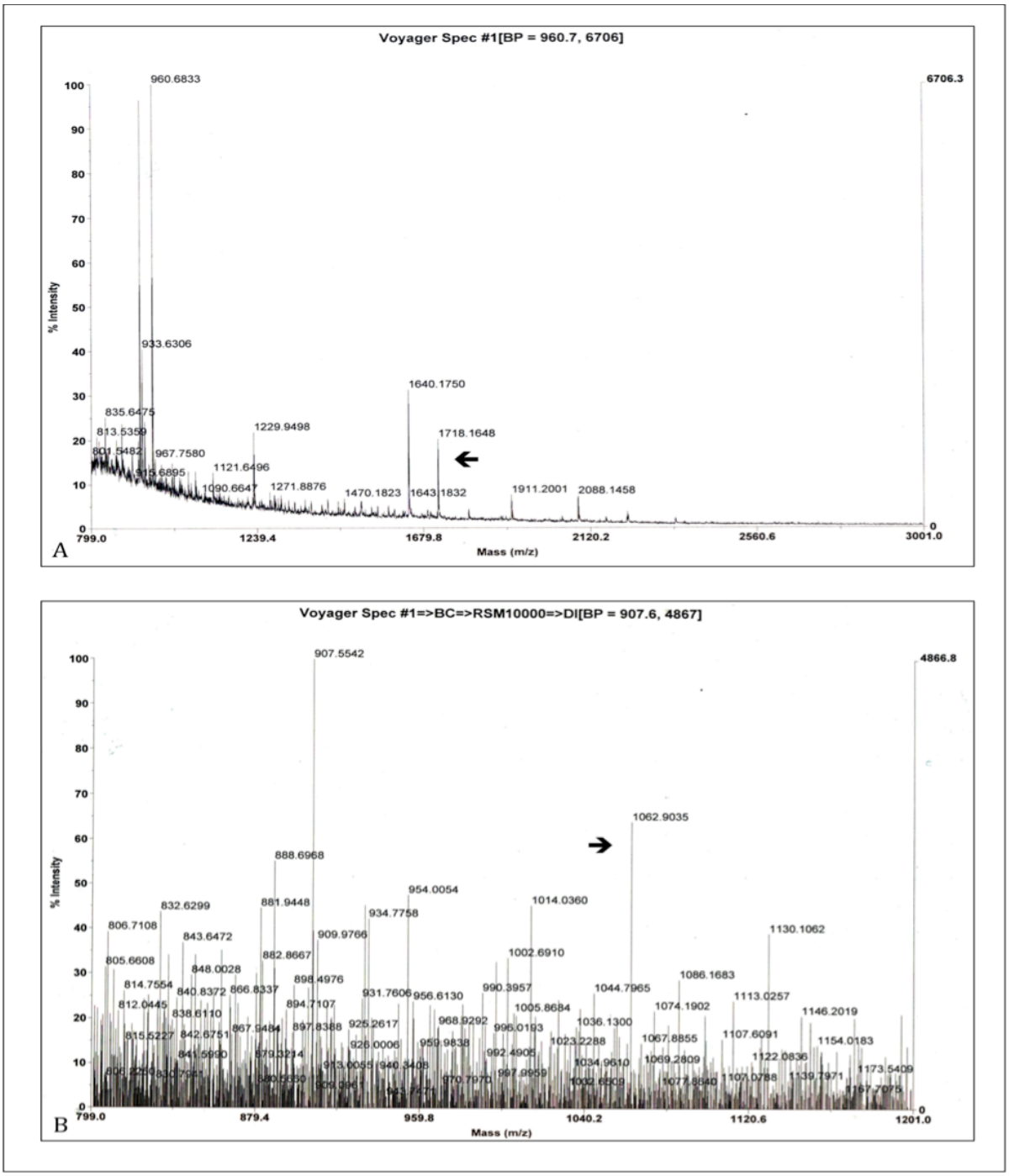

Figure 5.12. MALDI-TOF mass spectra of tryptic- and GluC-digested rGARP peptides that bound non-specifically to 'naked' goat anti-mouse IgG Dynabeads.

Panel A: "Sticky" peptides in trypsin digest. Panel B: "Sticky" peptides in GluC digest. GARP digests were incubated with 'naked' (non-antibody containing) magnetic Dynabeads. The Mass (m/z) range of GluC (panel B) was limited to 800-1200 instead of 800-3000 in order to be able to detect this peptide as there was another GluC-GARP peptide at 1790 [M+H⁺] that was saturating the signal, drowning out these minor peptides.

The signal intensity of the 1718 Da peptide enriched by mAb 2-D7 compared to its non-specific binding was 4.7×10^3 (Figure 5.11; panel A) and 1.47×10^3 (Figure 5.12; Panel A), respectively. Comparison of the peak intensities in the trypsin digest of GARP indicated that mAb 2-D7 bound 21 % of this peptide specifically, whereas the beads bound 6.6 % of this peptide non-specifically. This relatively large difference in peptide provided assurance that mAb 2-D7 specifically enriched this peptide. The signal intensity of peptide 1062 Da enriched by mAb 4-B7 compared to its non-specific binding was 1.5×10^4 (Figure 5.11; panel A) and 2.9×10^3 (Figure 5.12; panel B), respectively. Comparison of the peak intensities in the GluC digest of GARP indicated that mAb 4-B7 bound 31 % of this peptide specifically, while 5.9 % bound non-specifically. Once again this large difference gave assurance that mAb 4-B7 specifically enriched this peptide.

Glu-C cleaves peptide bonds C-terminally at glutamic acid (E) and with a 3000-fold lower rate at aspartic acid (Sorensen *et al.,* 1991), with its apparent specificity for glutamic acid higher in ammonium carbonate buffer, pH 7.8 (Houmard and Drapeau 1972). Trypsin cleaves peptide chains mainly at the carboxyl side of arginine (R) and lysine (Y) except when either is followed by proline (P) (Flannery *et al.,* 1989). Based on this information the 1718 Da mass enriched by mAb 2-D7 would be cleaved by GluC at E, ^{45}AL<u>E</u>-TASQSAVAAVVSSAR^{62}, as it contains an internal GluC cleavage site. Epitopes are usually approximately eight amino acids in length. Given that this peptide is destroyed by Glu-C cleavage, the epitope for mAb 2-D7 must be located towards the C-terminal end of this long sequence. When GARP was treated with trypsin the 1062 Da mass enriched by mAb 4-B7 was no longer detected; since the sequence contains an internal tryptic cleavage site (R) it was cleaved by trypsin at R: ^{141}SL<u>R</u>-LLATCE^{149}, thereby destroying the epitope. This

complementary outcome gave credence to the data.

Although I was working with a recombinantly expressed protein, I wanted to ensure that the predicted amino acid sequences of the peptides, based on the *in silico* digest, were accurate. To confirm the amino acid sequences of the epitope peptides, samples of interest were spotted onto a 384 well OptiT MALDI plate and the 1062 Da and 1718 Da peptide masses analyzed by MALDI-TOF-MS-MS on the 4800 reflector mass spectrometer with a 4000 explorer analyzer v3.5 installed at the UVic/Genome BC Proteomics Centre. The corresponding fragmentation spectra demonstrated that the predicted amino acid sequences were correct (Figure 5.13; panel A and B). DeNovo Explorer™ Software was used to interpret the high-energy collision-induced diffraction (CID) mass spectra of the 1062 Da and 1718 Da peptides acquired with a tandem time-of-flight MALDI mass spectrometer (Applied Biosystems 4800 Proteomics Analyzer). This program generates sequences that fit the peptide fragment ion data, as best it can, and these sequences are scored against the input peak list and ion intensities. The benefit to using this software is that it can be integrated with mass spectrometry-linked BLAST (MS BLAST) searching, thus allowing one to search the database for this specific sequence to find matches to the correct protein, and ultimately a high confidence score. However, this is not a 'perfect' program and there will be some amino acids that either aren't detected or are misidentified.

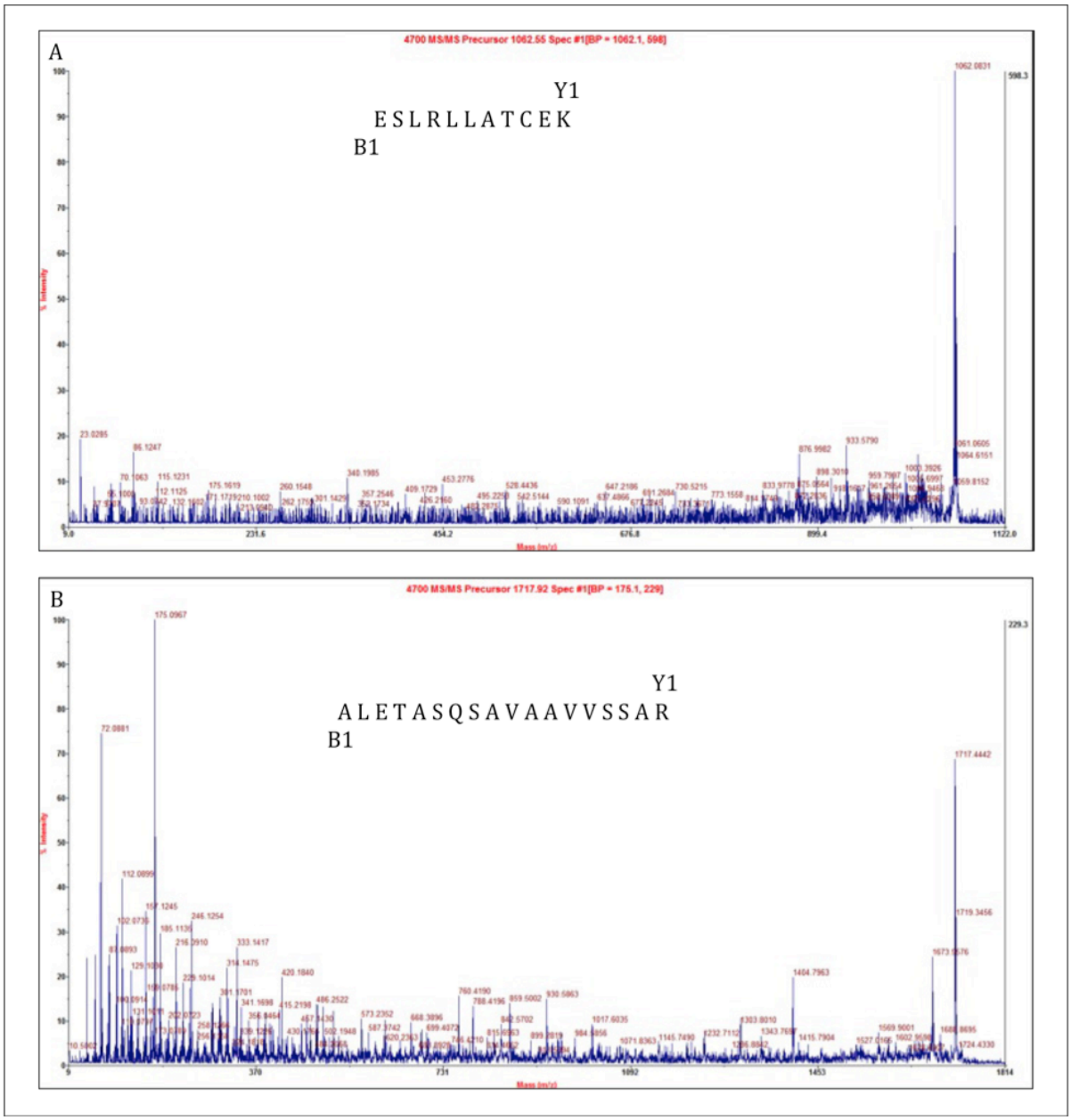

Figure 5.13. MS/MS mass spectra of epitope peptides, 1062 and 1718, enriched by immunoaffinity using mAb 4-B7 and 2-D7, respectively.

Panel A: MS/MS spectrum obtained for the 1062 peptide bound by mAb 4-B7. Panel B: MS/MS spectrum obtained for the 1718 peptide bound by mAb 2-D7. Indicated in both spectra are the amino acid sequences with the Y1 and B1 fragment ions labeled as starting points.

Based on the MS/MS masses of the y and b ions, the DeNovo software estimated the amino acid sequence of the 1718 Da peptide to be: (N-terminus)-SPETATNSAVLGAVSSAR-(C-terminus). Comparison of this predicted sequence to the sequence obtained from the *in silico* digest: (N-terminus)-ALETASQSAVAAVVSSAR-(C-terminus), identified eleven amino acids in common: (N-terminus)-**ETA**SAV***VVSSAR-(Cterminus); a 61 % identity. The DeNovo sequence was submitted to a BLAST search on NCBI, which correctly identified the peptide as belonging to GARP from *Trypanosoma congolense* with an E value of 7×10^{-4}. The amino acid sequence that DeNovo estimated for the 1062 Da peptide based on the MS/MS y and b ions was: (N-terminus)-SLRLLAM*NE-(C-terminus). Comparison to the sequence obtained from the *in silico* digest: (N-terminus)-SLRLLATCE-(C-terminus), identified seven amino acids in common: (N-terminus)-SLRLLA***E-(C-terminus); a 70 % identity. The DeNovo sequence was submitted to a BLAST search on NCBI, which correctly identified the peptide as belonging to GARP from *Trypanosoma congolense* with an E value of 0.17. A summary of the epitope mapping data is shown in Table 5.2.

Table 5.2. Summarized data: characterization of enriched trypsin/GluC rGARP peptides captured with either mAb 2-D7 or 4-B7 coupled Dynabeads.

	mAb 2-D7	**mAb 4-B7**
Digest sample	Trypsin	Glu-C
Enriched peptide mass [M+H]+	1718	1062
Peptide sequence[1]	^{45}ALETASQSAVAAVVSSAR^{62}	^{141}SLRLLATCE^{149}
Signal intensity of enriched peptide[2]	4.7×10^3	1.5×10^4
Signal intensity of peptide in digest[3]	2.22×10^4	4.9×10^4
Non-specific binding signal intensity[4]	1.47×10^3	2.9×10^3
% enrichment of peptide[5]	21 %	31 %
% binding non-specifically[6]	6.6%	5.9 %

[1] Based on predicted amino acid sequence from *in silico* digest, and confirmed by MS/MS.
[2] Intensity of the mAb-enriched specific peptide.
[3] Intensity of the specific peptide occurring in 10 µg of digested GARP.
[4] Intensity of the specific peptide binding non-specifically to 'naked' Dynabeads.
[5] Calculation: (signal intensity of peptide after enrichment ÷ signal intensity of the same peptide in the whole digest) × 100.
[6] Calculation: (signal intensity of peptide sticking non-specifically to beads ÷ signal intensity of the same peptide in the whole digest) × 100.

Previous analysis of mAb 2-D7 indicated that this antibody recognized GARP present in *T. congolense* as well as *T. simiae* CP 11, thereby being *Nannomonas*-specific (Figure 5.3; panel A) while mAb 4-B7 was found to be *T. congolense* savannah-specific (Figure 5.3; Panel B). With the epitope sequences now known I performed a multiple sequence alignment analysis of the full-length *GARP* genes of *T. congolense* 1/148 (Savannah), *T. congolense* K12 (Kilifi), and *T. simiae* CP 11 (Figure 5.14).

Figure 5.14. Multiple amino acid sequence alignment of the full length GARP genes of T. congolense Savannah 1/148 (accession L08055), T. simiae CP11 (accession AJ511258.1), and T. congolense Kilifi K12 (accession AF170298.1).

Dashes indicate gaps. The epitope-containing mAb 2-D7 and 4-B7 peptide sequences: ^{45}ALETASQSAVAAVVSSAR^{62}, and ^{141}SLRLLATCE^{149}, respectively, have been highlighted: green (T. simiae CP11), yellow (T. congolense 1/148) and red (T. congolense K12). (Expasy output- http://www.bioinformatics.nl/tools/muscle.html)

A closer look at the epitope sequences is shown in Table 5.3. The 2-D7 epitope regions are very similar between all three sequences. The *T. simiae* sequence, GVEAAIESAVAAATASAL, shares ten amino acids with the *T. congolense* sequences, ALETASQSAVAAVVSSAR, making it 55.5 % identical overall. Previous analysis by in-solution digestion of GARP with Glu-C resulted in the loss of this epitope, ^{45}ALE-TASQSAVAAVVSSAR^{62}, indicating that mAb 2-D7 likely bound to a region located towards the C-terminal end of this long sequence. If I assume that the epitope is approximately five to eight amino acids long, then the shared sequence SAVAA could possibly be the epitope. More work is required to test this possibility. While mAb 4-B7 was found to be *T. congolense* savannah-specific, its epitope sequence, when compared to that of *T. congolense*

K12 (Kilifi), is identical. Since the *GARP* gene sequence for the Kilifi strain, K45/1, that I used in my immunoblot analysis (Figure 5.3; panel B) has not been published it is possible that there may be enough amino acid dissimilarity that could result in this antibody not reacting to GARP. The Kilifi strain may also have a PTM that is concealing the epitope. This remains to be determined. No amino acid sequence similarity is shared between the *T. congolense* and *T. simiae* mAb 4-B7 epitope region. This is consistent with the immunoblot results where this mAb did not react to GARP present in PCF lysates of *T. simiae.* A summary of the epitopes of GARP of various trypanosomes is shown in Table 5.3

Table 5.3. Sequence comparison of the T. congolense* Savannah *GARP peptide sequences that contain the mAb 2-D7 and mAb 4-B7 epitope, with T. congolense Kilifi and T. simiae CP11 GARP sequences.

	2-D7 sequence[1]	**4-B7 sequence**
T. simiae CP 11	GVEAAIE**SAVAA**ATASAL	TLQQLAECD
T. congolense Savannah	ALETASQ**SAVAA**VVSSAR	SLRLLATCE
T. congolense Kilifi	ALETASQ**SAVTA**VVAFAR	SLRLLATCE

[1]*The potential mAb 2-D7 epitope are bolded.*

With the amino acid sequence of the epitopes for each anti-GARP mAb identified, their exact location on the tertiary structure of GARP was determined. Our collaboration with M. Boulanger's lab, which crystallized GARP and modeled its 3-D structure, made this possible. Figure 5.15 depicts the tertiary structure of GARP with the location of both epitopes indicated. The results of the epitope mapping on the tertiary structure are in perfect agreement with the immunofluorescence analyses on live and acetone-fixed trypanosomes.

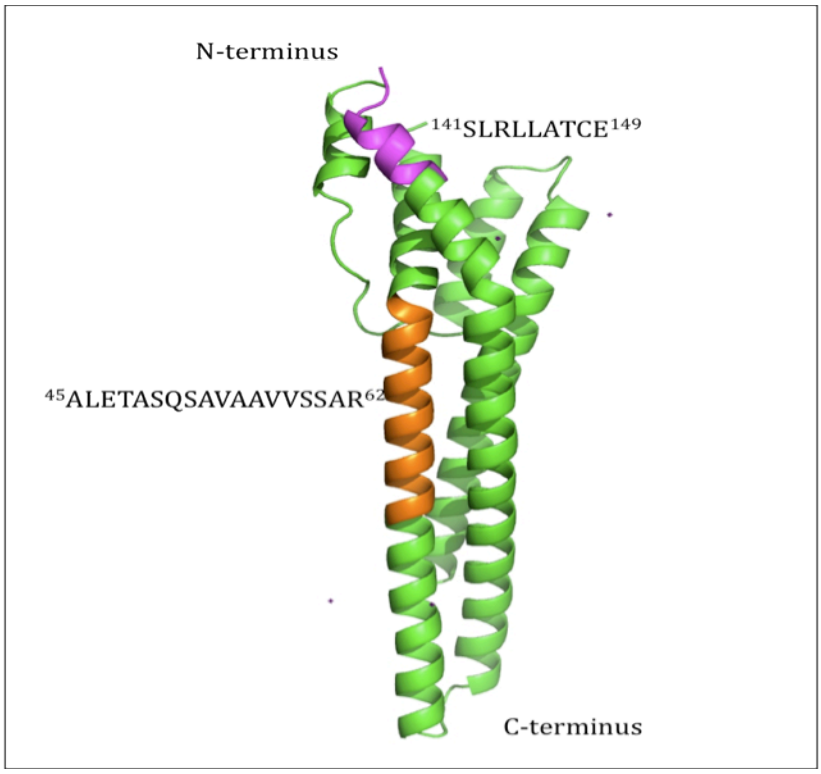

Figure 5.15. Three-dimensional model of crystallized GARP structure with mAb 2-D7 and 4-B7 epitopes mapped to their respective locations.
Knowing the amino acid sequences of the peptides selectively enriched by mAb 2-D7 and 4-B7 in the peptide capture assay, enabled localization of the mAb 2-D7 epitope, 1718 Da (orange) and the mAb 4-B7 epitope, 1062Da (purple).

The location of the epitope recognized by mAb 4-B7 (Figure 5.15; purple) at the N-terminal tip of GARP corresponds well with the surface immunofluorescence results (Figure 5.6) whereby mAb 4-B7 reacts strongly to GARP present on the surface of living trypanosomes. The location of the epitope recognized by mAb 2-D7 in an α helical region away from the surface disposed region correlates with the surface immunofluorescence results: mAb 2-D7 would not be able to bind to the surface of living trypanosomes since its epitope would be buried if the GARP molecules stack parallel to each other forming a surface coat. Some slight binding of this mAb was observed after acetone

fixation/permeabilization of trypanosomes however, indicating that some of the epitope became exposed by this treatment (Figure 5.6).

We found GARP interesting because of its differential expression, appearing early in the transformation of BSF to PF, during and after VSG displacement. The coexistence of VSG and GARP on the parasite surface implies some very interesting structural aspects to GARP that have been proposed but never studied at a structural level. With the Boulanger lab, we were now able to compare the 3-D structure of GARP with those of VSG1 and VSG2 (Figure 5.16), which are the only other trypanosome major surface molecules crystallized to date (Freymann *et al.,* 1990). All three proteins have in common a bent triple α-helix conformation. While the VSG1 and VSG2 have one short α-helix that extends into a random coil, the GARP helix is extended all the way up the molecule (Figure 5.16; blue). The similarity between all three protein structures demonstrates a core template that is indicative of function and higher order structure. It comes then as no surprise that VSG and GARP can be expressed simultaneously, as their similar structures would allow them to pack together quite tightly on the trypanosome membrane.

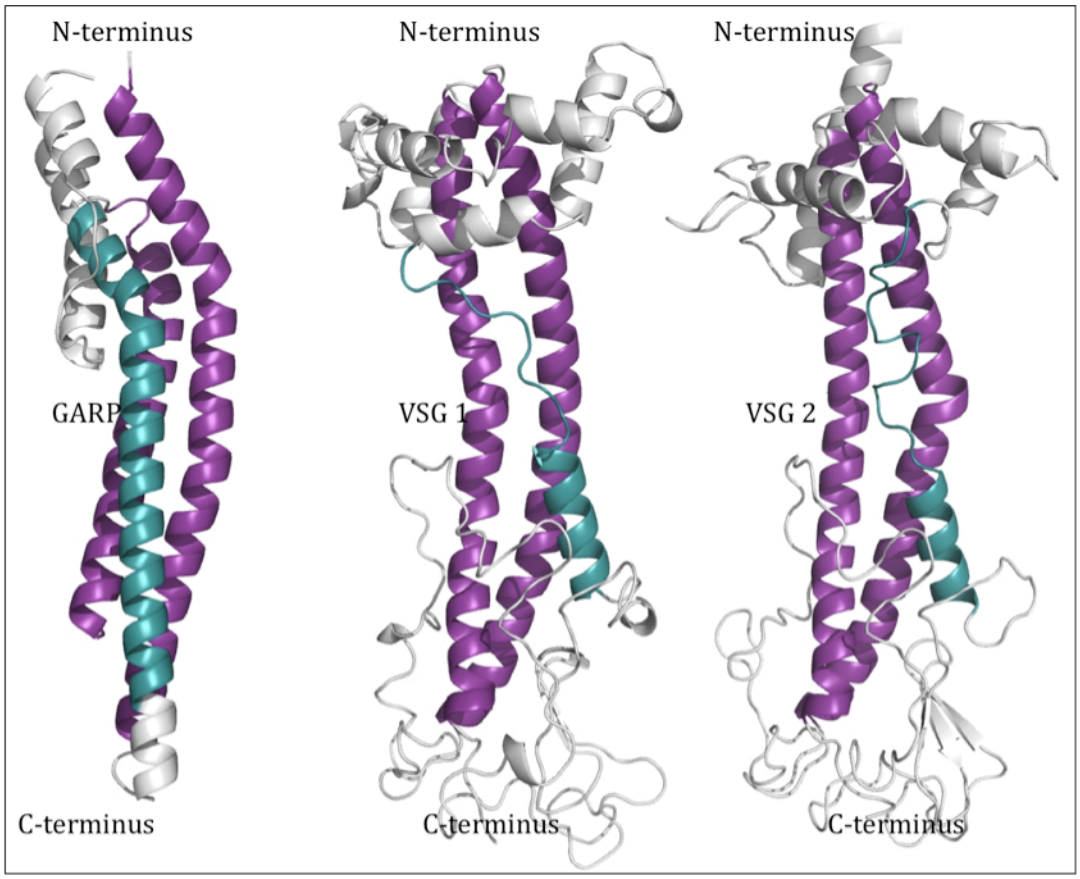

Figure 5.16. Three-dimensional model comparison of crystallized GARP, VSG1 and VSG2 structures to 1.65 Angstrom resolution.
Notice the slightly bent 3 α-helix bundle (purple and blue) that all three structures share in common. VSG1 and VSG2 'blue' helices are short, extending into a random coil, whereas the GARP 'blue' helix is extended all the way up the molecule.

5.4. Discussion

Much research has been done on GARP over the past several years. GARP, a relatively abundant surface molecule with a minimum copy number of 2×10^5 copies/cell, is expressed at low levels on early procyclic forms and at higher levels on late stage (established) procyclic forms and epimastigote forms. The GARP polypeptide is heavily glycosylated, immunodominant, highly charged, and subgenus *Nannomonas*-specific (Beecroft *et al.,* 1993; Bayne *et al.,* 1993; Thomson *et al.,* 2002). However, there still does

not exist an antibody probe specific to GARP that allows its detection, nor a crystal structure that might give clues as to its functional role. As part of my thesis research I set out to generate monoclonal antibodies to GARP in order to develop tools that would allow studies on GARP expression on trypanosomes throughout their life cycle. Such antibodies could also be used to map epitopes on the GARP polypeptide with the objective of studying the orientation of GARP on the trypanosome cell membrane. As my work progressed, so did the work in Dr. Martin Boulanger's lab to crystallize GARP. This allowed the localization of the epitopes on the 3-D structure of GARP, enabling us to postulate a general structure-function role for this surface molecule.

Two mAbs, 2-D7 and 4-B7, were generated by immunizing mice with rGARP and chosen from among several, for their individual characteristics. MAb 2-D7 recognizes the GARP polypeptide in parasites of the subgenus *Nannomonas (T. congolense* and *T. simiae)* but does not bind to cell surface epitopes. MAb 4-B7 recognizes GARP polypeptide in *T. congolense* Savannah strains only, and binds to surface-disposed epitopes.

Years ago the theory was put forth that GARP may play an active role in tropism of trypanosomes of the subgenus *Nannomonas* in tsetse, for example in migration from the midgut to the proboscis or mouthparts, where trypanosomes of this subgenus differentiate into mammal-infective metacyclic forms (Roditi *et al.,* 1989; Beecroft *et al.,* 1993; Bayne *et al.,* 1993; Garside *et al.,* 1995). Although the *GARP* gene was originally described only in *T. congolense* (Beecroft *et al.,* 1993; Bayne *et al.,* 1993) in 2000 *GARP* gene homologues were found in *T. simiae* and *T. godgreyi* (Asbeck *et al.,* 2000). It is therefore, very encouraging that I have generated a mAb, 2-D7, that reacts in immunoblots to proteins of *T. congolense* and *T. simiae* but not *T. brucei.* This is the first demonstration of a GARP polypeptide in *T. simiae*,

providing further evidence to support the aforementioned theory that GARP may be important to development of subgenus *Nannomonas* trypanosomes within the fly.

Another theory proposed soon after the discovery of procyclic surface molecule EP (Richardson *et al.,* 1988), is that in addition to GARP playing an active role in tropism, it plays a role in protection of the parasite membrane from digestive enzymes, antibodies and complement in the tsetse midgut (Roditi and Pearson 1990; Stebeck and Pearson 1994). With the elucidation of the GARP crystal structure in the Boulanger lab, it became clear that GARP had an overall similar shape to VSG molecules, the only other trypanosome surface membrane protein crystallized to date (Freymann *et al.,* 1984 and 1990). This allowed us to theorize that VSG and GARP could possibly form a surface coat by packing together on the trypanosome during VSG-loss and GARP-gain that occurs during differentiation from BSF to procyclic forms in the tsetse midgut. In such a way the parasite membrane would always remain covered, and thus protected from active complement factors present in the bloodmeal. The abundance and regularly spaced glutamic acids and alanines of GARP has been suggested to allow the formation of an extended α-helical shape to the protein (Beecroft *et al.,* 1993). This same formation was modeled for the EP repeat procyclins of *T. brucei* (Roditi *et al.,* 1989) after nuclear magnetic resonance determined that the EP dipeptide repeats confer a highly extended, rodlike structure upon procyclin (Evans *et al.,* 1986). In addition, immunogold electron microscopy showed GARP to be densely packed on the cell surface of *T. congolense* (Bayne *et al.,* 1993). The crystal structure of GARP was found to be in perfect agreement with both of these earlier observations: it is made up of a triple-helix bundle that, like VSG, has a slightly bent conformation, with an extended α-

helical rod-like structure indicative of a molecule that can pack tightly together forming a dense coat (Freyman *et al.,* 1984; Metcalf *et al.,* 1987).

It has been documented that while GARP is primarily expressed on the surface of late-stage PF and EMFs, there is a brief weak expression on early PF shortly after entering the midgut (Bütikofer *et al.,* 2002; Utz *et al.,* 2006). Immunofluorescence analysis, using mAb 4-B7, on parasites taken from different compartments of infected tsetse during the course of an infection with *T. congolense* also showed GARP to be weakly expressed by early midgut forms ten days post-infection, as well as later forms found in the mouthparts that were determined by DAPI, to be properly differentiated EMFs.

The two anti-GARP mAbs chosen were interesting for three reasons. First, one was surface reactive and the other not, indicating that while for one the epitope is exposed on the surface of living cells, the other is either concealed by its location and/or disposition on the membrane, or it is obscured by some sort of PTM on the live parasites. Second, determining the location of their respective epitopes would give us a clue as to the orientation of GARP on the surface of the parasites, and thirdly having a mAb that reacts with molecules in both *T. congolense* and *T. simiae* enables us to compare the amino acid sequences of the GARP polypeptides to confirm the epitope. The epitope sequences for mAb 2-D7 and 4-B7 were determined to be ^{45}ALETASQSAVAAVVSSAR^{62}, and ^{141}SLRLLATCE^{149}, respectively. According to the GARP 3-D structure, the sequences for mAb 2-D7 and 4-B7 are located along one of the three alpha helical regions that comprise the helical bundle and at the most N-terminus region of GARP, respectively. Epitopes at these positions help explain my immunofluorescence results. The peptide epitope sequence recognized by mAb

2-D7, located on a region of GARP that is part of a triple-helix bundle, would be inaccessible to antibodies since this is a region of GARP that would allow close packing of individual GARP molecules, forming a shield on the parasite surface, thereby concealing this epitope-containing sequence on live trypanosomes.

A comparative analysis of the mAb 2-D7 epitope-containing peptide, ^{45}ALETASQSAVAAVVSSAR^{62}, in the GARP sequences from *T. congolense* Savannah, *T. congolense* Kilifi and *T. simiae* revealed a five amino acid stretch, SAVAA, that was shared. The epitope for mAb 4-B7, which is nine amino acids long, is not shared between *T. congolense* Savannah and *T. simiae*, consistent with the finding that mAb 4-B7 is *T. (Nannomonas) congolense* Savannah-specific.

The antibodies used by other researchers to detect GARP have, until now, either been polyclonal, thus recognizing several different GARP epitopes, or reactive towards carbohydrate epitopes (mAb 491) that are shared among several surface glycoconjugates expressed in *T. congolense*. To continue the study of GARP more specific probes are required. I have successfully generated two, renewable, monoclonal antibodies, which react specifically to the GARP polypeptide, and have determined the location of their cognate epitopes on the GARP molecule. These reagents were shown to be useful probes for monitoring the expression of GARP in different species and life cycle stages of trypanosomes, including those taken directly from tsetse and will be of value for studies on structure-function relationships of these interesting surface molecules.

Chapter 6. Conclusions and Potential Future Research.

African trypanosomes are complex protozoan parasites that are digenetic, evasive and adaptive. These parasites prepare themselves for survival in the mammalian host or tsetse vector by expressing a number of different stage-specific GPI-anchored surface carbohydrates, proteins or glycoproteins. Of particular interest are those molecules expressed on the surface of trypanosomes that occupy the tsetse vector. While several major surface molecules have been identified, it is not known whether or not they function in protection against proteases (or antibodies or complement), facilitate signaling, or are involved in differentiation, transit or tropism in the fly. Research has been hindered in part by the lack of sufficient probes that can be used to conclusively identify, locate and perhaps interfere with these proteins and their function. The identification of antibody epitopes and their characterization is extremely important to aid this problem. Understanding antibody specificity at the molecular level provides the key to optimizing their use as research tools. It was a major goal of my thesis to characterize, in depth, a selection of monoclonal antibodies to the cell surface molecules of *T. congolense:* HRP, CESP and GARP.

Throughout this thesis I have described the successful generation and characterization of five monoclonal antibodies: one mAb is specific to the TcHRP repeat, EPGENGT, and two mAbs each are specific to the CESP and GARP polypeptides. I used these to track the expression of the three surface molecules in different species and life cycle stages of trypanosomes, including those taken directly from infected tsetse. I was not so lucky in identifying the epitope of the two anti-CESP mAbs due to the complications associated with expressing, purifying and crystallizing recombinant proteins. I was

successful in identifying the epitopes for both anti-GARP mAbs, as well as localizing their positions on the 3-D crystal structure of GARP. I believe these probes will be very useful tools to study all three *T. congolense* surface molecules and to aid in determining their functional roles.

No functional role has been assigned to any of the cell surface molecules expressed in insect forms of African trypanosomes. However, it is clear that future studies of these molecules will be aided by the antibody probes generated during my thesis research and it is my hope that the next generation of researchers will put them to good use. Looking ahead, there are several possibilities for future research projects, both short term and long term. First, it should be a priority to complete the epitope identification for the two anti-CESP mAbs, and to obtain a crystal structure of CESP. Having a crystal structure will help elucidate whether or not CESP is an adhesion molecule, as has been postulated (Sakurai *et al.,* 2008). In addition, antibody inhibition studies, gene knockouts or transcript knockdowns by RNA interference might shed light on the role of CESP in adhesion to tsetse mouthparts and/or whether or not it is required for metacyclogenesis of EMFs into MCFs. The anti-CESP mAbs will be invaluable for such research.

Of no less importance, studies on the role of GARP and TcHRP can be approached in similar ways. All *Glossina* sp. harbour and provide a stable host environment to *Wigglesworthia glossinidius*, and *Sodalis glossinidius,* primary and beneficial obligate, Gram-negative bacterial endosymbionts. One method for determining the possible function of CESP, GARP and TcHRP in the tsetse vector would be to genetically transform these symbionts to produce single chain antibodies of the surface reactive mAbs. This would

allow expression of mAbs in the vector such that these antibodies could bind to parasite molecules. If mAb binding inhibits or interferes with parasite differentiation or tropism and/or metacyclogenesis in such paratransgenetically-altered tsetse, then useful targets for interference with trypanosome transmission will have been identified.

All three of the major surface molecules of *T. congolense* that I studied in my research are abundant and temporally expressed in trypanosomes during their transit through the tsetse vector and thus almost surely play important functional roles. Understanding these roles will allow development of strategies for interference with transmission of trypanosomes and may, in the future, contribute to the control of African trypanosomiasis.

References

Acosta-Serrano, A., Cole, R.N., Mehlert, A., Lee, M.G., Ferguson, M.A., and Englund, P.T. 1999. The procyclin repetoire of *Trypanosoma brucei*. Identification and structural characterization of the Glu-Pro-rich polypeptides. *J. Biol. Chem.* 274: 29763-29771.

Acosta-Serrano, A., Vassela, E., Liniger, M., Kunz Renggli, C., Brun, R., Roditi, I., and Englund, P.T. 2001. The surface coat of procyclic *Trypanosoma brucei:* programmed expression and proteolytic cleavage of procyclin in the tsetse fly. *Proceedings of the National Academy of Sciences of the United States of America.* 98: 1513-1518.

Anderson, N.G., and Anderson, N.L. 1979. *Molecular Anatomy*. Behring Inst. Mitt, 63: 169-210.

Asbeck, K., Kurath, U., Roditi, I., and Gibson, W. 2004. *Trypanosoma (Nannomonas) simiae* and *T. (N.) godreyi* have genes encoding glutamic acid and alanine-rich proteins. *Mol. Biochem. Parasitol*. 134; 159-162.

Asbeck, K., Ruepp, S., Roditi, I., and Gibson, W. 2000. GARP is highly conserved among *Trypanosoma congolense* Savannah. *Mol. Biochem. Parasitol*. 106: 303-6.

Bayne, R.A.L., Kilbride, E.A., Lainson, A., Tetley, L., Barry, D. 1993. A major surface antigen of procyclic stage *Trypanosoma congolense*. *Mol. Biochem. Parasitol*. 61: 295-310.

Beattie, P. and Gull, K. 1997. Cytoskeletal architecture and components involved in the attachment of *Trypanosoma congolense* epimastigotes. *Parasitology*. 115: 47-55.

Beecroft, R.P., Roditi, I., Pearson, T.W. 1993. Identification and characterization of an acidic major surface glycoprotein from procyclic stage *Trypanosoma congolense*. *Mol. Biochem. Parasitol*. 61: 285-94.

Bienen, E. J., Webster, P., Fish, W. R. 1991. *Trypanosoma (Nannomonas) congolense*: changes in respiratory metabolism during the life cycle*. Exp Parasitol.* 73: 403-412.

Blum M.L., Down J.A., Gurnett A.M., Carrington M., Turner M.J., Wiley D. 1993. A structural motif in the variant surface glycoproteins of *Trypanosoma brucei*. *Nature*. 362:603-609.

Broden, A. 1904. Les infections à trypanosomes au Congo chez l'homme et les animaux (communication préliminaire). *Bull. Soc. Étud. Colon*. 11, 116-139.

Bruce, D. Harvey, D., Hamerton, A.E., Davey, J.B. and Bruce, M. 1912. The morphology of *T. simiae,* sp. nov. *Proc. Biol. Sci*. 85, 477-481.

Budd, L.T. 1999. DFID-Funded Tsetse and Trypanosomiasis Research and Development Since 1980 (Economic Analysis). Department for International Development.

Bütikofer, P., Ruepp, S., Boschung, M., and Roditi, I. 1997. 'GPEET' procyclin is the major surface protein of procyclic culture forms of *Trypanosoma brucei brucei* strain 427. *Biochem. J.* 326: 415-423.

Bütikofer, P., Vassella, E., Boschung, M., Kunz Renggli, C., Brun, R., Pearson, T.W., and Roditi, I. 2002. Glycosylphosphatidylinositol-anchored surface molecules of *Trypanosoma congolense* insect forms are developmentally regulated in the tsetse fly. Mol. *Biochem. Parasitol*. 119: 7–16.

Bütikofer, P., Vassella, E., Ruepp, S., Boschung, M., Civenni, M., Seebeck, T., *et al.* 1999. Phosphorylation of a major GPI-anchored surface protein of *Trypanosoma brucei brucei* during transport to the plasma membrane. *J. Cell Sci*. 112: 1785-95.

Chandra, M., Liniger, M., Tetley, L., Roditi, I., and Barry, J.D. 2004. TsetseEP, a gut protein from the tsetse *Glossina morsitans*, is related to a major surface glycoprotein of trypanosomes transmitted by the fly and to the products of a *Drosophila* gene family. *Insect Biochem. And Molec Biol*. 34: 1163–1173.

Chattopadhyay, A., Jones, N., Nietlispach, D., Nielsen, P., Voorheis, H., Mott, H., Carrington, M. 2005. Structure of the C-terminal Domain from *Trypanosoma brucei* Variant Surface Glycoprotein MITat1.2. *J Biol Chem*. 280: 7227-7235.

Chen, X., Juhasz, P., Papayannopoulos, I.A., Parker, K.C., Regelson, M.E., Spencer, D., and Zhu, X. 2003. A Computer Program Combining De Novo Sequencing of High Energy CID Mass Spectra of Peptides with MS BLAST Searching for Protein Identification. Applied Biosystems, Framingham, MA; 2Paracel Inc., Pasadena, CA.

Clausen, P. H., Sidibe, I., Kabore, I. and Bauer, B. 1992. Development of multiple drug resistance of *Trypanosoma congolense* in Zebu cattle under high natural tsetse fly challenge in the pastoral zone of Samorogouan, Burkina Faso. *Acta Tropica.* 5: 229-236.

Clayton, C.E., and Mowatt, M.R. 1989. The procyclic acidic repetitive proteins of *Trypanosoma brucei.* Purification and post-translational modification. *J. Biol. Chem.* 264: 15088-15093.

Colmerauer, M.E.M., Davis, C.E., and Pearson, T.W. 1989. The trypanosome surface glycoprotein procyclin is expressed only on tsetse-fly vector stages of the parasite. *Parasitol. Res*. 76: 171-3.

Cross, G.A.M. and Manning, J.C. 1973. Cultivation of *Trypansoma brucei* spp. in semi-defined and defined media. *Parasitol*. 67: 315-331.

De Raadt, P. 1999. The history of sleeping sickness. In: Gilles HM, ed. *Protozoal diseases*. London, Arnold.

Evans, D.E., Ellis, D.S., Stamford, S. 1979. Ultrastructural studies of certain aspects of the development of *Trypanosoma congolense* in *Glossina morsitans morsitans. J Protosool*. 26: 557-63.

Field, M.C., Menon, A.K., and Cross, G.A. 1991. A glycosylphosphatidyl-inositol protein anchor from procyclic stage *Trypanosoma brucei*: lipid structure and biosynthesis. *EMBO J.* 10: 2731-2739.

Fiennes, R.N.T.-W. 1970. Pathogenesis and pathology of animal trypanosomiasis. In: H.W. Mulligan, Editor, *The African Trypanosomiases*, Allen and Unwin, London. 729–750.

Fish, W.R., Muriuki, C.W., Muthiani, A.M., Grab, D.G. and Lonsdale-Eccles, J.D. 1989. Disulfide bond involvement in the maintenance of the cryptic nature of the cross-reacting determinant of metacyclic forms of *Trypanosoma congolense. Biochem*. 28: 5415-5421.

Flannery, A.V., Beynon, R.J., and Bond, J.S. 1989. Proteolysis of Proteins for Sequencing Analysis and Peptide Mapping. *IRL Press, Oxford, U.K.* Proteolytic Enzymes: A Practical Approach. R.J. Beynon and J.S. Bond, eds.

Freymann, D., Down, J., Carrington, M., Roditi, I., Turner, M., and Wiley, D. 1990. 2·9Å resolution structure of the N-terminal domain of a variant surface glycoprotein from *Trypanosoma brucei. J. Mol. Biol.* 216: 141-160.

Freymann, D., Metcalf, M., Turner, J., and Wiley, D. 1984. 6Å resolution X-ray structure of a variable surface glycoprotein from *Trypanosoma brucei. Nature (Lond).* 311: 167-169.

Garside, L.H and Gibson, W.C. 1995. Absence of the glutamic acid alanine-rich protein (GARP) genes in the *Nannomonas* species of *Trypanosoma simiae* and *T. godfreyi. Mol. Biochem. Parasitol*. 74: 211-5.

Gibson, W., and Bailey, M. 2003. The development of *Trypanosoma brucei* within the tsetse fly midgut observed using green fluorescent trypansomes. *Kinetoplastid Biol. Dis.* 2: 1-13.

Gibson, W.C., Stevens, J.R., Mwendia, C.M.T., Makumi, J.N., Ngotho, J.M., Ndung'u, J.M. 2001. Unravelling the phylogenetic relationships of African trypanosomes of suids. *Parsitol*. 122: 625-31.

Gray M.A., Cunningham, I., Gardiner, P.R., Taylor, A.M., and Luckins, A.G. 1981. Cultivation of infective forms of *Trypanosoma congolense* from trypanosomes in the proboscis of *Glossina morsitans. Parasitol.* 82: 81-95

Greganova, E., Bütikofer, P., and Acosta-Serrano, A. 2010. The protease resistant surface (PRS) glycoconjugate from *Trypanosoma congolense* has an inositol-acylated glycosylphophatidylinositol anchor, containing a significant proportion of myristate at the *sn*-2 position. *Mol. Biochem. Parasitol.* 171: 50-54.

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

Haines, L.R., Lehane, S.M., Pearson, T.W., and Lehane, M.J. 2010. Tsetse EP protein protects the fly midgut from trypanosome establishment. *PLoS Pathog*. 6: e1000793.

Helm, J.R., Hertz-Fowler, C., Aslett, M., Berriman, M., Sanders, M., Quail, M.A., Soares, M.B., Bonaldo, M.F., Sakurai, T., Inoue, N., Donelson, J.E. 2009. Analysis of expressed sequence tags from the four main developmental stages of *T. congolense. Mol Biochem Parasitol*. 168: 34-42.

Hendry, K.A.K. and Vickerman, K. 1988. The requirement for epimastigote attachment during division and mtacyclogensis in *Trypanosoma congolense. Parasitol Res*. 74: 1-19.

Hide G. 1999. History of sleeping sickness in East Africa. *Clin Microbiol Rev*. 12: 112-125.

Hirumi H., and Hirumi K. 1991. *In vitro* cultivation of *Trypanosoma congolense* bloodstream forms in the absence of feeder cell layers. *Parasitol.* 102: 225-236.

Hoare, C.A. 1970. Systematic Description of the Mammalian Trypanosomes of Africa, in H. Mulligan & W. Potts: *The African Trypanosomiases*. London, UK: George Allen and Unwin Ltd.. ISBN 0-04-614001-8.

Hoffman, R. 1954. Zur Fortpflanzungsbiologie und zur intrauterinen Entwicklung von *Glosina palpalis. Acta Tropica.* 11: 1-57.

Houmard, J., and Drapeau, G.R. 1972. Staphylococcal Protease: A Proteolytic Enzyme Specific for Glutamoyl Bonds. *Proc. Natl. Acad. Sci. USA*. 69: 3506- 3509.

Janssen, J.A.. and Wijers, D.J. 1974. *Trypanosoma simiae* at the Kenya coast. A correlation between virulence and the transmitting species of Glossina. *Annual Tropical Medicine Parasitol.* 68: 5-19.

Jemmerson, R. and Patterson, Y. 1986. Mapping antigenic sites on proteins: implications for the designe of synthetic vaccines. *BioTechniques.* 4: 18-31.

Jemmerson, R., and Patterson, Y. 1986. Mapping epitopes on a protein antigen by the proteolysis of antigen-antibody complexes. *Science.* 232: 1001-1004.

Jordan, A.M. 1976. Tsetse flies as vectors of trypanosomes. *Veterinary Parsitol.* 2: 143-152.

Kristjanson, P.M., Swallow, B.M., Rowlands, G.J., Kruska, R.L and de Leeuw, R.N. 1999. Measuring the costs of African animal trypanosomiasis, the potential benefits of control and returns to research. *Agr. Syst*. 59: 79-98.

Laemmli, U.K. 1970. Cleavage of structural proteins duing the assembly of the head of bacteriophage T4. *Nature*. 227: 680-5.

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

Majiwa, P.A.O. and Webster, P. 1987. A repetitive deoxyribonucleic acid sequence distinguishes *Trypanosoma simiae* from *T. congolense. Parasitol*. 95: 543-558.

Mattioli, R., and Slingenbergh, J. 2008. Programme Against African Trypanosomiasis (PAAT). Retrieved April 24 2008, from http://www.fao.org/ag/againfo/programmes /en/paat/home.html.

Metcalf, P., Blum, M., Freymann, D., Turner, M., and Wiley, D.C. 1987. Two variant surface glycoproteins of *Trypanosoma brucei* of different sequence classes have similar 6Å resolution X-ray structures. *Nature (Lond.).* 325: 84-86.

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

Mihok, S., Zweygarth, E., Munyoki, E.N., Wambua, J., Kock, R. 1994. *Trypanosoma simiae* in the white rhinoceros (*Ceratotherium simum)* and the dromedary camel *(Camelus dromedarius)*. *Veterinary Parasitol.* 53: 191-6.

Moloo, S.K., Zweygarth, E., Sabwa, C. L. 1992. Virulence of *Trypanosoma simiae* in pigs infected by *Glossina brevipalpis, G. pallidipes* or *G. morsitans centralis. Annual Trop Med Parasitol.* 86: 681-3.

Molyneux, D.H., Lavin, D.R., and Elce, B. 1979. A possible relationship between Salivarian trypanosomes and Glossina labrum mechano-receptors. *Ann Trop Med Parasitol*. 73: 287-90.

Mookherjee, N. and Pearson, T.W. 2001. Surface molecules of procyclic forms of *Trypanosoma simiae* and *Trypanosoma congolense*, members of the subgenus *Nannomonas*, share immunodominant carbohydrate epitopes. *Mol. Biochem. Parasitol*. 118: 123-6.

Mookherjee, N. and Pearson, T.W. 2002. *Trypanosoma simiae* and *Trypanosoma congolense*: surface glycoconjugates of procyclic forms-the same coats on different hangeres? *Exp. Parasitol.* 100: 257-68.

Mowatt, M.R., and Clayton, C.E. 1987. Developmental regulation of a novel repetitive protein of Trypanosoma brucei. *Mol. Cell. Biochem*. 7: 2838-2844.

Nagamune, K., Acosta-Serrano, A., Uemura, H, *et al.* 2004. Surface sialic acids taken from the host allow trypanosome survival in tsetse fly vectors. *J Exp Med.* 199: 1445-50.

Neuhoff, V., Arold, N., Taube, D. and Ehrhardt, W. 1988. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis*. 9: 255-262.

Oli, M.W., Cotlin, L. F., Shiflett, A.M. and Hajduk, S.L. 2006. Serum resistance-associated protein blocks lysosomal targeting of trypanosome lytic factor in *Trypanosoma brucei*. *Parasitol. Res*. 80: 177-181.

Parham, P. 1983. On the fragmentation of monoclonal IgG1, IgG2a, and IgG2b from BALB/c mice. *J. Immunol.* 131: 2895-2902.

Pays, E. and Nolan, D.P. 1998. Expression and function of surface proteins in *Trypanosoma brucei. Mol. Biochem. Parasitol.* 91: 3-36.

Pearson, T.W., Moloo, S.K., Jenni, L. 1987. Culture form and tsetse fly midgut form procyclic *Trypanosoma brucei* express common proteins. *Mol. Biochem. Parasitol*. 25: 273-8.

Porter, R. R. 1959. The hydolysis of rabbit γ-globulin and antibodies with crystalline papain. *Biochem. J.* 73: 119-126.

Rangarajan, D., Harvey, T.I., and Barry, J.D. 2000. Characterization of the loci encoding the glutamic acid and alanine rich protein of *Trypanosoma congolense. Mol. Biochem. Parasitol.* 105: 281-290.

Raska, C.S., Parker, C.E., Sunnarborg, S.W., Pope, R.M., Lee, D.C., Glish, G.L., and Borchers, C.H. 2003. Rapid and sensitive identification of epitope-containing peptides by direct matrix-assisted laster desorption/ionization tandem mass spectrometry of peptides affinity-bound to antibody beads. J. *Am. Soc. Mass Spectrom.* 14: 1076-1085.

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

Richardson, J.P., Jenni, L., Beecroft, R.P., and Pearson, T.W. 1986. Procyclic tsetse fly midgut forms and culture forms of African trypanosomes share stage- and species-specific surface antigens identified by monoclonal antibodies. *J. Immunology.* 136: 2259-2264.

Robinson, D.R., and Gull, K. 1991. Basal body movement as a mechanism for mitochondrial genome segregation in the trypanosome cell cycle. *Nature* (Lond.). 352: 731-734.

Roditi, I., Carrington, M., Turner, M. 1987. Expression of a polypeptide containing a dipeptide repeat is confined to the insect stage of *Trypanosoma brucei*. *Nature*. 325: 272-4.

Roditi, I., Furger, A., Ruepp, S., Schurch, N. and Bütikofer, P. 1998. Unravelling the procyclin coat of *Trypanosoma brucei. Mol. Biochem. Parasitol.* 91: 117-130.

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

Roditi, I and Pearson, T.W. 1990. The procyclin coat of African trypanosomes. *Parasitology Today.* 6: 79-81.

Roditi, I.H., Schwartz, H., Pearson, T.W., Beecroft, R.P., Liu M.K., Richardson, J.P., *et al.* 1989. Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of *Trypanosoma brucei*. *J. Cell Biol*. 108: 737-46.

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

Saleh, K.M., M.J.B. Vreysen, S.S., Kassim, F.W., Suleiman, K.G., Juma, Z.-R., Zhu, H., Pan, V.A., Dyck, and U. Feld- mann. 1997. The successful application of the sterilein-sect technique (SIT) for the eradication of *Glossina austeni* (*Diptera: Glossinidae*) from Unguja island (Zanzibar), pp. 438-445. In Twenty Fourth Meeting of the International Scientifc Council for Trypanosomosis Research and Control, Maputo, Mozambique, 1997. OAU/IBAR, Nairobi, Kenya.

Sakurai, T., Sugimoto, C., Inoue, N. 2008. Identification and molecular characterization of a novel stage-specific surface protein of *Trypanosoma congolense* epimastigotes. *Mol. Biochem. Parasitol*. 161: 1-11.

Sheshberadaran, H. and Payne, L.G. 1988. Protein antigen-monoclonal antibody contact sites investigate by limited proteolysis of monoclonal antibody-bound antigen: protein "footprinting." *Proc. Natl. Acad. Sci. USA.* 85: 1-5.

Smith, D.H., Pepin, J., and Stich, A.H.R. 1998. Human African trypanosomiasis: an emerging public health crisis. *British Medical Bulletin.* 54: 341-355.

Sørensen, S.B., Sørensen, T.L. and Breddam, K. 1991. Fragmentation of proteins by S. aureus strain V8 protease. Ammonium bicarbonate strongly inhibits the enzyme but does not improve the selectivity for glutamic acid. *FEBS*. 294: 195–197.

Stebeck, C.R., and Pearson, T.W. 1994. Major Surface Glycoproteins of Procyclic Stage African Trypanosomes. *Exp. Parasitol.* 78: 432-436.

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.

Thevenaz, P., and Hecker, H. 1980. Distribution and attachment of *Trypanosoma* (*Nannomonas*) *congolense* in proximal part of the proboscis of *Glossina morsitans morsitans*. *Acta Trop*. 37: 163-75.

Thomson, L.M., Lamont, D.J., Mehlert, A., *et al*. 2002. Partial structure of Glutamic Acid and Alanine-rich Protein, a Major Surface Glycoprotein of the Insect Stages of *Trypansoma congolense*. *J. Biol. Chem*. 277: 48899-48904.

Tiralongo, E., Schrader, S., Lange, H., Lemke, H., Tiralongo, J., and Schauer, R. 2003. Two *trans*-sialidase forms with different sialic acid transfer and sialidase activities from *Trypanosoma congolense, J Biol Chem.* 278: 23301–23310.

Treumann, A., Zitzmann, N., Hülsmeier, A., Prescott, A.R., Almond, A., Sheehan, J., *et al.* 1997. Structural characterization of two forms of procyclic acidic repetitive protein expressed by procyclic forms of *Trypanosoma brucei. J Mol. Biol.* 269: 529-47.

Utz, S., Roditi, I., Kunz Renggli, C., Almeida, I.C., Acosta-Serrano, A., and Bütikofer, P. 2006. *Trypanosoma congolense* procyclins: unmasking cryptic major surface glycoproteins in procyclic forms. *Eukaryotic Cell.* 5: 1430-1440.

Vassella, E., Acosta-Serrano, A., Studer, E., Lee, S.H., Englund, P.T. and Roditi, I. 2001. Multiple procyclin isoforms are expressed differentially during the development of insect forms of *Trypanosoma brucei. J. Mol. Biol.* 312: 597-607.

Vassella, E. van den Abbeele, J., Bütikofer, P., Kunz Renggli, C., Furger, A., Brun, R., *et al.* 2000. A major surface glycoprotein of *Trypanosoma brucei* is expressed transiently during development and can be regulated post-transcriptionally by glycerol or hypoxia. *Genes Deve.* 14: 615-26.

Vickerman, K. 1962. The mechanism of cyclical development in trypanosomes of the Trypanosoma brucei sub-group: an hypothesis based on ultrastructural observations. *Transactions of the Royal Society of Tropical Medicin and Hygiene*. 56: 487-95.

Vickerman, K. 1965. Polymorphism and mitochondrial activity in sleeping sickness trypanosomes. *Nature.* 208: 762-6.

Vickerman, K. 1966. Genetic systems in unicellular animals. *Science Progress*. 54: 13-26.

Vickerman, K. 1969. On the surface coat and flagellar adhesion in trypanosomes. *Journal of Cell Science*. 5: 163-93.

Vickerman, K. 1969. The fine structure of Trypanosoma congolense in its bloodstream phase. *Br. Med. Bull.* 41: 105-114.

Vickerman, K. 1985. Developmental cycyles and biology in pathogenic trypanosomes.. *Journal of Protozoology.* 16: 54-69.

Vickerman, K., Tetley, L., Hendry, K.A., Turner, C.M. 1988. Biology of African trypanosomes in the tsetse fly. *Biol Cell*. 109-119.

Welburn , S.C., and Maudlin, I. 1999. Tsetse-trypanosome interactions: rites of passage. *Parasitology Today.* 15: 399-403.

Welburn, S.C., Picozzi, K., *et al.* 2001. Identification of human infective trypanosomes in animal reservoirs of sleeping sickness in Uganda by means of serum-resistance-associated (SRA) gene. *The Lancet.* 358: 2017-2019.

WHO Media centre. 2006. *Fact sheet N°259: African trypanosomiasis or sleeping sickness*. http://www.who.int/mediacentre/factsheets/fs259/en/).

Young, C. J. and Godfrey, D.G. 1983. Enzyme polymorphism and the distribution of *Trypanosoma congolense* isolates. *Annals of Tropical Med. and Parasitol.* 77: 467-481.

Zweygarth, E. and Rotcher, D. 1987. The occurance of *Trypanosoma (Nannomonas) simiae in the cerebrospinal fluid (CS;) of domestic pigs. Parasitol Res.* 73: 479-80.

Appendix 2. Abbreviations

1-D	*one dimensional*
A	*alanine*
A°	*angstrom*
AAT	*animal African trypanosomiasis*
ACTH	*adrenocorticotropic hormone*
BC	*British Columbia*
BLAST	*basic local alignment search tool*
bp	*base pair*
BSF	*bloodstream form*
C	*cysteine*
°C	*celsius*
cDNA	*complimentary deoxyribonucleic acid*
CESP	*congolense epimastigote-specific surface protein*
CHO	*carbohydrate*
CID	*collision-induced diffraction*
DAPI	*4',6-Diamidino-2-phenylindole*
DMEM	*Dulbecco's modified eagles medium*
DTT	*Dithiothreitol*
E	*glutamic acid*
EP	*glutamic acid-proline*
EPGENGT	*glutamic acid-proline-glycine-glutamic acid-aspartic acid-glycine-threonine*
ELISA	*enzyme-linked immunosorbant assay*
EMF	*epimastigote form*
FACS	*fluorescence activated cell sorter*
FBS	*fetal bovine serum*
FITC	*fluorescein isothiocyanate*
Gal	*galactose*
GARP	*glutamic acid/ alanine-rich protein*
GC-MS	*gas chromatography-mass spectrometry*

Glc	*glucose*
GlcN	*glucosamine*
GlcNAc	*N-acetylglucosamine*
GPEET	*glycine-proline-glutamic acid-glutamic acid-threonine*
GPI	*glycosyl-phosphatidyl-inositol*
GST	*glutathioine S-transferase*
HAT	*human African trypanosomiasis*
HAT	hypoxanthine, aminopterin, thymidine
H+L	heavy and light
ht	*human transferrin*
HT	*hypoxanthine, thymidine*
Ig	*immunoglobulin*
IL-6	*interleukin-6*
IP	*intraperitoneally*
IV	*intravenously*
K	*lysine*
kDa	*kilodalton*
kDNA	*kinetoplast deoxyribonucleic acid*
km	*kilometer*
µ	*micro*
M	*molar*
mAb(s)	*monoclonal antibodies*
MALDI	*matrix-assisted laser desorption*
MALDI-TOF-TOF	*tandem mass spectrometry*
Man	*Mannose*
MCF	*metacyclic culture form*
MEM	*modified eagle's medium*
MF	*metacyclic form*
µg	*microgram*
µL	*microlitre*
µm	*micrometer*

µM	*micromolar*
mg	*milligram*
min	*minute*
mL	*millilitre*
mM	*millimolar*
mRNA	*messanger RNA*
MS-MS	*tandem mass spectrometry*
NCBI	*national center for biotechnology information*
N-X-S/T	*asparagine-X-serine/threonine*
OPI	*oxaloacetate, pyruvate, insulin.*
PBS	*phosphate buffered saline*
PCF	*procyclic culture form*
PCR	*polymerase chain reaction*
PEG	*polyethylene glycol*
PF	*procyclic form*
rpm	*revolutions per minute*
PRS	*protease-resistant surface molecule*
PSG	*phosphate buffered saline + 1% glucose*
PTM	*post-translational modification*
PV	*proventriculus*
PVDF	*polyvinylidene difluoride*
Q	*glutamine*
R	*arginine*
rCESP	*recombinant congolense epimastigote-specific protein*
rGARP	*recombinant glutamic acid/alanine-rich protein*
RNA	*ribonuecleic acid*
S	*serine*
SA	*sialic acid*
SDS-PAGE	*sodium dodecyl sulphate-polyacrylamide gel electrophoresis*
T	*threonine*
T .b. brucei	*Trypanosoma brucei brucei*

TC	*tissue culture*
TcHRP	*T. congolense heptapeptide repeat protein*
Thr	*threonine*
TOF	*time of flight*
Tris	*tris(hydroxymethyl)aminomethane*
TWP	*Terry W. Pearson*
UK	*United Kingdom*
USA	*United States of America*
V	*valine*
VSG	*variant surface glyocprotein*
v/v	*volume per volume*
WHO	*world health organization*
w/v	*weight per volume*
Y	*tyrosine*