IDENTIFICATION AND CHARACTERIZATION OF THE KINETOPLASTID MEMBRANE PROTEIN-11 (KMP-11) GENE LOCUS FROM *TRYPANOSOMA BRUCEI* AND ITS GENETIC DELETION

by

MICHAEL ADRIAN ANTHONY BRIDGE
B.Sc., University of Victoria, 1996

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of
Biochemistry and Microbiology

We accept this dissertation as conforming
to the required standard

[redacted]

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

[redacted]

Dr. Robert Olafson, Department Member
(Department of Biochemistry and Microbiology)

[redacted]

Dr. William Hintz, Outside Member
(Department of Biology)

[redacted]

Dr. Robert Burke, External Examiner
(Department of Biology)

University of Victoria

Supervisor: Dr. Terry W. Pearson

ABSTRACT

The kinetoplastid membrane protein (KMP)-11 is an amphipathic membrane associated molecule that was first identified in *Leishmania donovani* and was subsequently found to have a ubiquitous distribution among kinetoplastid parasites, including African trypanosomes. KMP-11 has been identified, by immunofluorescence, along the length of the flagellum and in the flagellar pocket. A 3500 bp genomic DNA fragment containing the KMP-11 sequence was identified, isolated and subcloned from a P1 bacteriophage library containing *T.brucei* strain TREU 927/4 genomic DNA. A comprehensive library for automated DNA sequencing was produced by nebulizing the 3500 bp P1 fragment containing the KMP-11 locus. Seventy-two individual sequencing reads averaging 550bp per read were aligned and gave an 11.5 fold sequence redundancy. This allowed accurate assembly of the highly repetitive KMP-11 sequence. The KMP-11 locus contains a tightly packed array of four highly conserved KMP-11 repeats found within a 2200 bp region. The repeated genes contain only two positions at which the sequences diverge. In both cases transitions have occurred in the third position of the codon and do not alter the predicted amino acid sequence of the protein. The four genes are arranged in two sets of larger repeats that are separated by an intergenic region of 209 bp. This intergenic region is the only non-duplicated portion of the entire KMP-11 locus and was the key to assembling the sequence. Upstream of each of the four genes is a conserved 5' untranslated region containing long poly-pyrimidine rich elements and putative splice acceptor sites that may play an important role in accurate splicing of the KMP-11 polycistronic transcript.

Once the KMP-11 locus had been fully characterized, gene deletion studies were pursued in the hope of elucidating the function of KMP-11. The deletion cassettes

contained genomic flanking sequences from the KMP-11 locus surrounding either the neomycin or phleomycin drug resistance gene. The first round of knockouts produced the expected null/+ genotype having replaced one of the two KMP-11 loci with the neomycin drug resistance deletion cassette. Subsequent attempts to produce complete null/null mutants, under a wide variety of conditions, have been unsuccessful and suggest that the full deletion of KMP-11 in *Trypanosoma brucei* is lethal.

Examiners:

[redacted]

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

[redacted]

Dr. Robert Olafson, Department Member
(Department of Biochemistry and Microbiology)

[redacted]

Dr. William Hintz, Outside Member
(Department of Biology)

[redacted]

Dr. Robert Burke, External Examiner
(Department of Biology)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS USED:

ADP, adenosine diphosphate
AMP, adenosine monophosphate
ATP, adenosine triphosphate
BSF, blood stream form
BLAST, Basic Local Alignment Search Tool
CO_2, carbon dioxide
CsCl, cesium chloride
dATP, 2'-deoxyadenosine triphosphate
dCTP, deoxycytidine triphosphate
ddATP, dideoxyadenosine triphosphate
ddCTP, dideoxycytidine triphosphate
ddGTP, dideoxyguanine triphosphate
ddNTP, dideoxynucleotide triphosphate
ddTTP, dideoxythymine triphosphate
dGTP, 2'-deoxyguanine triphosphate
dNTP, deoxynucleotide triphosphate
dTTP, deoxythymine triphosphate
DNA, deoxyribonucleic acid
EDTA, ethylene-diaminetetra-acetic acid
ELISA, enzyme linked immunosorbent assay
EP repeat, Glutamic acid-Proline repeat
ESAG, expression site associated gene
FBS, fetal bovine serum
FITC, fluorescein-isothiocyanate
GPI, glycosylphosphatidylinositol
GPEET repeat, glycine-proline-$(\text{glutamic acid})_2$-threonine repeat
GrESAG, gene related to expression site associated gene
HDL, high density lipoprotein
HPLC, high performance liquid chromatography
IFN, interferon
IPTG, isopropyl-b-D-galactosidase
ISG, invarient surface glycoprotein
KMP-11, kinetoplastid membrane protein – 11
LB, Luria-Bertani

LDL, low density lipoprotein
LPG, lipophosphoglycan
mAb, monoclonal antibody
MEM, minimal essential media
mRNA, messenger ribonucleic acid
Neo, neomycin
PAG, procyclin associated gene
PARP, procyclin acidic repetitive protein
PBS, phosphate-buffered saline
pBS, pBluescript
PCF, procyclic culture form
PCR, polymerase chain reaction
PEG, polyethelene glycol
Phleo, phleomycin
pI, isoelectric point
PSSA-2, procyclic stage surface antigen-2
RNA, ribonucleic acid
SDS-PAGE, sodium dodecylsulfate polymerase gel electrophoresis
Sh ble, phleomycin drug resistance gene
TFA, trifluoroacetic acid
Tht, trypanosome hexose transporter
TLF, trypanosome lytic factor
TNF, tumor necrosis factor
UTR, untranslated region
VSG, variant surface antigen
WHO, world health organization
ZM, Zimmerman's media

ACKNOWLEDGEMENTS

Thanks to Dr. Ben Koop and Dr. Qun Zhou (Department of Biology, University of Victoria, Victoria, British Columbia, Canada) for their collaborative efforts in automated sequencing and generation of sequencing libraries, Dr. Isabel Roditi and Dr. Stefan Reupp (Department of General Microbiology, University of Bern, Bern, Switzerland) for their expert assistance in parasite transformation and selection techniques and their hospitality during my visit to Bern, Dr. W. Robert McMaster (Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada) for informative discussions about gene knockout studies, Dr. Robert Burke (Department of Biology, University of Victoria, Victoria, British Columbia, Canada) for his adept assistance with confocal immunofluorescence microscopy, Dr. William Hintz (Department of Biology, University of Victoria, Victoria, British Columbia, Canada) for his advice about "long" polymerase chain reactions, Dr. Francis Choy (Department of Biology, University of Victoria, Victoria, British Columbia, Canada) for his positive personal interest in my work; and Dr. Fran Nano (Department of Biochemistry, University of Victoria, Victoria, British Columbia, Canada) for providing use of his electroporation apparatus.

I also thank Robert Beecroft, Neeloffer Mookherjee and Jennifer Chase, the scientists and friends who have generated a fun and supportive environment in the lab.

Special thanks must go to my supervisory committee, who have been insightful, supportive and generous with their knowledge. Your positive attitudes and insight have made me proud to be a fellow scientist.

Dr. Olafson, thank you for your insight and supportive personal commentary, I am ready to run with the *it*.

I am a grateful recipient of Scholarships, Fellowships and research grants from various groups. Thanks to the Arthritis Society of British Columbia (Victoria Chapter) for the Ray Hadfield Memorial Fellowship, to the Natural Sciences and Engineering Research Council of Canada for a post graduate scholarship (PGS A), and to the University of Victoria for the President's Research Scholarship, and the Dr. Julius Schleicher Graduate Scholarship.

To my mom, dad and sisters, you have made the most profound difference in my life with your unwavering love and support. You have helped me realize that all of my dreams are attainable. Jeffie, many thanks for all the time, effort, and love you have given me, it has bolstered my sprit and made it a joy to come home each day.

Terry, your kindness, good humor, desire to impart knowledge about lab work and about life, have made this period of study a wonderful learning experience. You are a special person that I will always value you as a mentor and a friend.

GENERAL INTRODUCTION

1.1 African Trypanosomoses

Diseases caused by parasites are perhaps the most socially and economically important health problems in the world. Of these, the African trypanosomoses have severe and broad reaching effects on humans and animals. The World Health Organization [WHO] considers African trypanosomoses to be one of the most serious threats to human health and a serious obstacle to the development of agriculture in Africa [WHO, 1979]. The eukaryotic protozoan parasites which causes the disease, African trypanosomes, are carried from host to host via the tsetse fly and thus the disease is largely contained within the natural tsetse fly infested regions of Africa. The most heavily studied of the African trypanosomes include the causative agents of human sleeping sickness (*Trypanosoma brucei gambiense* and *Trypanosoma brucei rhodesiense*) and those effecting domestic animals (*Trypanosoma vivax, Trypanosoma congolense, Trypanosoma simiae, Trypanosoma brucei brucei* and *Trypanosoma suis*) all of which are prevalent in tsetse infected areas within Africa [Mulligan *et al.*, 1970]. In recent years upwards of 20,000 new cases of human trypanosomoses have been reported each year, and an estimated 50 million people are at risk of catching either the West African-Gambian sleeping sickness or the East African-Rhodesian sleeping sickness [WHO, 1979]. In 1998, there are reports of an epidemic in Sudan, Central Africa Republic and Zaire, affecting more than 300,000 people [Zimmer, 1998]. The Rhodesian form of sleeping sickness caused by *T.b.rhodesiense* is a more acute disease usually leading to death within a year, while the Gambian form caused by *T.b.gambiense* has a more chronic course [Greenwood *et al.*, 1972]. After enduring fluctuating waves of parasitemia, untreated victims of African sleeping sickness often succumb to the disease when the parasite invades the central nervous system and brain, causing menigoencephalitis [Shapiro *et al.*, 1986]. Patients may also die of coma-induced starvation, or of secondary infections such as pneumonia due to immune suppression. Human suffering and economic hardship is perhaps more widely felt from Nagana, the cattle equivalent of African sleeping sickness, which greatly reduces

livestock production throughout Africa. As many as four million square miles of sub-Saharan Africa cannot be used for cattle raising, which inevitable results in poverty and malnutrition within the endemic areas. The livestock pathogens, *T.congolense*, *T.vivax* and *T.b.brucei* produce anemia, cachexia (wasting), heart damage and immune suppression in the cattle. This suppression of immune responses causes death via secondary infections such as pneumonia [Shapiro *et al.*, 1986].

1.2 The Classification of African Trypanosomes

African trypanosomes are unicellular eukaryotic protozoan parasites that possess a kinetoplast, a dense organelle containing mini and maxi circles of mitochondrial DNA, located near the base of the flagellum [Vickerman, 1976]. These parasites have a digenetic life cycle in which they live alternatively in the blood stream of vertebrate hosts and the gut of arthropod vectors. African trypanosomes belong to the genus Trypanosoma; family Trypanosomatidae; suborder Trypanosomatina; order Kinetoplastida; class Zoomastigophorea; subphylum Mastigophora; phylum Sarcomastigophora; subkingdom protozoa; kingdom Protista. Organisms from the genus Trypanosoma are divided into 2 sections, that of Stercoraria and Salivaria (Figure 1). These sections differ primarily in the location at which the parasite develops within the vectors [Molyneux and Ashford, 1983]. *T.cruzi*, the causative agent of Chagas' disease is representative of the Stercoraria section which finish development in the terminal gut of the vector and are transmitted through its feces during feeding. The parasite once deposited on the skin of the mammalian host gains access through epidermal lesions or mucus membranes. Alternatively, African trypanosomes and other species of the Salivaria section complete development in the anterior portion of the vector's digestive tract and are transmitted through the saliva during feeding [Pearson *et al.*, 1993; Haag *et al.*, 1998].

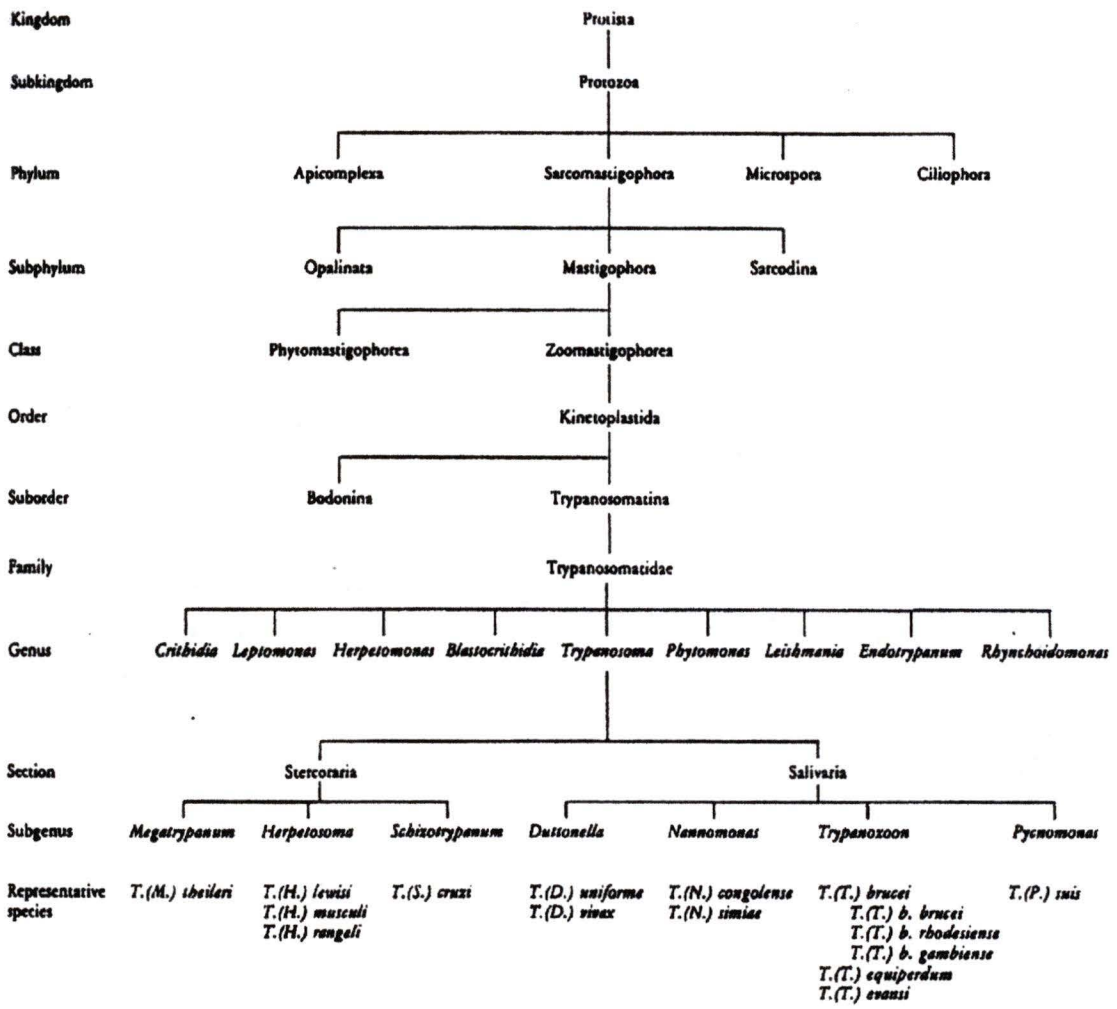


Figure 1: The systematics of African trypanosomes. Adapted from Molyneux and Ashford, 1983.

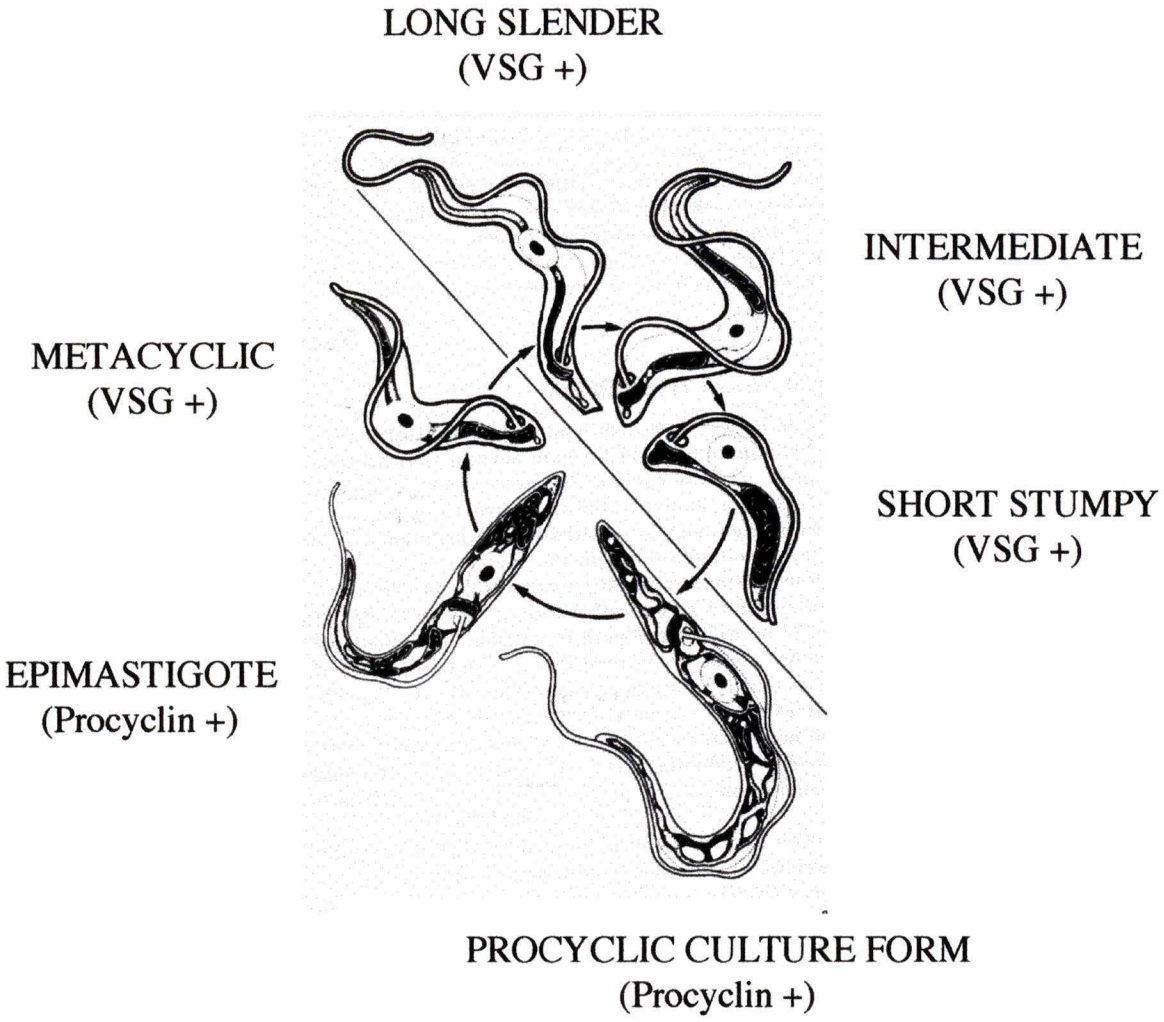
1.3 The Trypanosome Life Cycle

As digenetic parasites that live alternatively in the bloodstream of host vertebrates and the midgut of tsetse fly vectors, the African trypanosomes must adapt to a diverse set of conditions in vastly different environments. To accommodate the separate environments, the trypanosome life cycle has developed to contain forms that are morphologically, physiologically, and antigenically distinct from one another. The parasite life cycle is also characterized by a succession of proliferative stages associated with establishing infection and non-proliferative stages which serve to aid in transmission of the parasite [Vanhamme and Pays, 1995]. The life cycle of *Trypanosoma brucei*, which has been elucidated in detail, will be discussed further as a representative model (Figure 2).

The long-slender bloodstream form (BSF) trypanosomes, found in the mammalian host, are actively proliferating cells that have repressed mitochondrial function [Vickerman, 1965]. This inefficient metabolic state is adopted by the BSF since the host bloodstream provides nutrients. As the rudimentary mitochondrion does not allow oxidative phosphorylation, these organisms metabolize glucose to pyruvate through the glycolytic pathway in specialized organelles termed glycosomes [Vanhamme and Pays, 1995]. The BSF are covered with a homogeneous layer of the variant surface glycoprotein (VSG). The VSG coat of the trypanosome population continuously changes throughout the term of the infection, allowing evasion of host immune defenses. A series of parasitemic waves is observed during infection, where each successive wave of trypanosomes presents a novel, uniform VSG barrier to the immune system. During each wave, while the antibody response of the infected organism eliminates most of the trypanosomes, individual parasites expressing a new VSG coat are able to proliferate and repopulate the host [Donelson *et al.*, 1998]. These antigenically distinct trypanosomes generate the subsequent parasitemic wave and the characteristic pattern seen during antigenic variation continues.

As the infection progresses, a number of the long slender parasites differentiate into short stumpy BSF forms that possess semi-developed mitochondria [Vickerman, 1965]. These trypanosomes have ceased to divide and are pre-adapted for transmission into the

IN THE MAMALIAN HOST



IN THE TSETSE FLY

Figure 2: The life cycle of *T. brucei* spp. in the tsetse fly vector and mammalian host. Adapted from Vickerman *et al* ,1980.

Fly, where more efficient metabolic processes are required in the relatively carbohydrate deprived environment of the tsetse fly midgut [Shapiro *et al.*, 1984 and 1986]. Once taken up by the fly, the mitochondrion of the short stumpy BSF becomes elaborated with cristae as the cells differentiate into the actively dividing procyclic forms. The procyclic trypanosomes adapt to life in the tsetse fly midgut where they respire on proline, using the Krebs cycle in aerobic respiration as a major source of energy. These procyclic forms that were once thought to be uncoated, after the loss of their VSG, are covered with another glycoprotein coat termed procyclin. This procyclin coat is comprised of several forms of the procyclin molecules, including forms with extensive EP repeats and GPEET repeats [Roditi *et al.*, 1998].

The procyclic forms migrate from the fly midgut to the fly salivary glands where they differentiate into epimastigotes [Vickerman, 1962] which have the unique attribute of having their nucleus posterior to the kinetoplast (Figure 2). The epimastigotes then differentiate into non-dividing metacyclic trypanosomes which are pre-adapted for transmission, having re-acquired the VSG coat so essential for parasite survival in the mammalian host [Shapiro *et al.*, 1986].

Once the metacyclic trypanosomes are injected back into the mammalian host during tsetse fly feeding, they differentiate into the rapidly dividing long slender BSFs and bring the host to vector life cycle of the African trypanosome full turn. Integral to the life cycle and the successful adaptation of the trypanosomes to its disparate environments are the major surface antigens, the procyclins and VSG. These antigens will be discussed in detail later in the introduction.

The different life cycle stages of African trypanosomes require differential expression of genes. It is therefore of interest to understand several aspects of gene expression in order to understand the biology of the parasites.

1.5 Control of Gene Expression in Trypanosomes

1.5.1 Polycistronic Transciptional Unit

Protein coding genes in trypanosomes are generally arranged as clusters of tandem repeats of two or more similar genes [Clayton *et al.*, 1992; Amar et al., 1988; Bridge *et al.*, 1998] and lack class II introns [Pays *et al.*, 93]. Related genes, such as those encoding alpha and beta tubulin [Imboden *et al.*, 1987] or unrelated genes such as in the aldolase locus [Vijayasarathy *et al.*, 1990] may be grouped together or interspersed between the gene repeats. The repeated genes within the cluster as well as the functional intergenic regions are co-transcribed as a long polycistronic precursor RNA [Clayton, 1992]. Multiple monocistronic mRNAs are subsequently generated through the addition of a 5' mini exon and a 3' polyadenylation tail [Schurch *et al.*, 1994]. The individual units of a multigene cluster may have very different expression profiles even though they are transcribed together in a polycistronic fashion. This implies that the regulation of gene expression may be controlled at the posttranscriptional level and, in fact, evidence suggests that some of the information for this regulation may be encoded in the intergenic regions that were once thought to have little function [Vassella *et al.*, 1994].

1.5.2 Processing of the Polycistronic Unit

When polycistronic transcripts are processed into smaller functional mRNA units, intergenic sequences are excised, and both a spliced leader sequence and a poly A tail are added. The spliced leader, sometimes termed the mini-exon, is a 39 nucleotide sequence that is found in every mature mRNA and is encoded by mini-exon donor DNA [Lee, 1997]. The splice acceptor site for trypanosomes, as for higher eukaryotes, is localized to an AG dinucleotide that is preceded by polypyrimidine tracts [Schurch *et al.*, 1994; Vassella *et al.*, 1994]. While processed trypanosome mRNA contains a poly A tail, the common hexamer signal sequence for polyadenylation found in other eukaryotes is not present. Trypanosomes do not contain an absolute consensus site for polyadenylation and

it is now thought that both trans-splicing and polyadenylation may be linked and occur within a defined region [Layden *et al.*, 1988]. Deletion studies have shown that the polyadenylation site of the upstream gene and the splice site of the following gene has a conserved distance of between 100 to 150 nucleotides [Hug *et al.*, 1994]. This is consistent with the hypothesis that the trans splicing of a 5' mini exon and the addition of a 3' poly A tail are linked through the binding of processing machinery to polypyrimidine tracts present in the intergenic region [Vanhamme *et al.*, 1995]. An illustration of this mechanism is presented in Figure 3. Although there is no absolute consensus for polyadenylation, recent studies have outlined four elements that may play a role in accurate polyadenylation [Schurch *et al.*, 1994]. Of primary importance, are small stretches of 3 to 6 adenosine residues found at the poly A site itself. An intervening region of 60 to 90 bp usually follows this site. A pyrimidine rich stretch is then observed after which a trinucleotide YAG is present, and finally a short distance downstream there is often a second pyrimidine rich region followed by one or two more copies of the YAG trinucleotide [Schurch *et al.*, 1994]. The polypyrimidine tracts are an important element in both the splicing and polyadenylation of the polycistronic transcripts and this motif likely acts as the bridge that links these two processes coordinating the production of functional mRNA molecules.

1.6 RNA Editing

RNA editing was initially characterized in the mitochodrial DNA of trypanosomes [Benne *et al.*, 1986], but similar processes have been identified in a wide variety of organisms including humans (mitochondrial DNA) [Ojala *et al.*, 1981] and plants (mitochondrial and chloroplast DNA) [Cattaneo *et al.*, 1992]. The mitochondrial DNA of trypanosomes is found in a unique organelle termed the kinetoplast, as a complex organization of concatenated mini and maxi circles. The majority of mitochondrial proteins characterized thus far have incomplete encoding genes and thus RNA editing generates functional

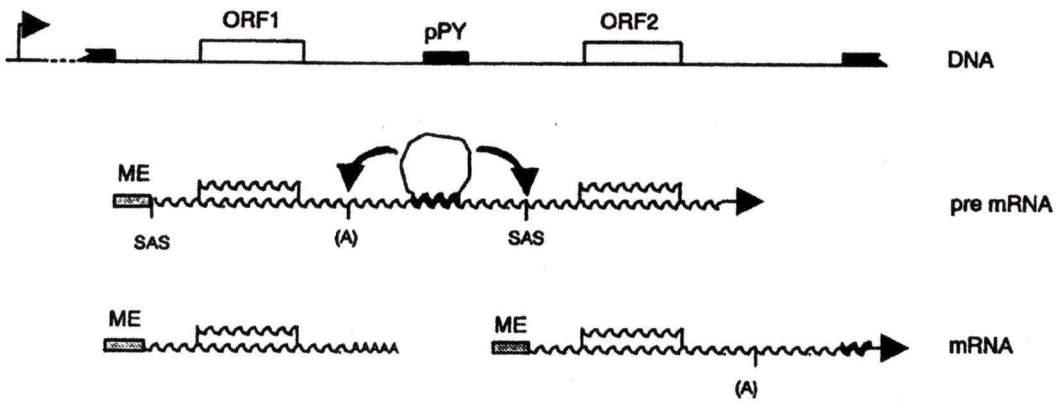


Figure 3: Illustration of the coupling of trans-splicing and polyadenylation. The following abbreviations are used: ORF, open reading frame; pPY, pol pyrimidine tracts; ME, mini-exon; SAS, splice acceptor site; (A), polyadenylation site. Adapted from Vanhamme and Pays, 1995.

mRNAs. This process most commonly involves the insertion of 1 to 13 Uridylate (U) residues but can also involve deletion of up to 6 U residues at a given site within the premature mRNA [Benne *et al.*, 1994]. The editing occurs posttranscriptionally using small guide RNA molecules that are encoded in the mini and maxi mitochondrial DNA circles. These guide RNA molecules which are between 55 to 70 nucleotides in length, provide all the editing information required to generate functional mRNA from incomplete genes [Benne, 1994]. Through this process, open reading frames that resemble pseudogenes encoding frame shifts or missing translation initiation codons are edited in such a fashion that functional proteins are produced. As well as its short term attributes, RNA editing has been hypothesized to play the more fundamental role of generating sequence diversity by producing increased protein heterogeneity thus acting as a source of rapid genetic variation [Landweber and Gilbert, 1993].

1.7 Major Surface Glycoproteins of African Trypanosomes.

1.7.1 Variant Surface Glycoproteins

The variant surface glycoprotein (VSG) covers the surface of the BSF trypanosomes with up to 1 X 10^7 molecules per trypanosome. This molecule associates as a dimer adopting an extended configuration perpendicular to the membrane. The relatively conserved C-terminus of VSG (~ 75 amino acids) is attached to the plasma membrane via a phospholipase C sensitive GPI anchor and to the highly variable N-terminus (~350 amino acids) region [Blum *et al.*, 1993]. All the VSG proteins are N-glycosylated near or within the C-terminal region and have a relatively similar tertiary structure even with the highly variable N-terminal domain. The VSG molecules are packed very tightly forming a protective coat over the entire parasite. The conformation of the molecule is such that only a small stretch of amino acids from the N-terminus is exposed at the surface and the C-terminus is entirely buried [Pays and Nolan, 1998]. This type of organization shields potentially vulnerable invariant proteins present on the membrane, prevents recognition of

the conserved C-terminus epitopes on the VSG molecule and safeguards the trypanosome against lytic elements such as complement, antibodies and IFN gamma, found in the BSF environment [Donelson *et al.*, 1998]. VSG is highly immunogenic and is detected and removed by the host immune system. In fact, there is continual removal of the majority of trypanosomes carrying the major antigenic coat, which are inevitably replaced by parasites with a newly expressed VSG [Shapiro *et al.*, 1986]. While the immune system stops acute killing of the host from uncontrolled overgrowth of the parasite, the VSG antigenic variation facilitates a chronic infection in which a continuous series of parasitemic waves is observed.

Antigenic variation, which occurs spontaneously at 10^{-2} switches per cell per generation, is a function of the genetic organization of the VSG genes. More than 2% of the genomic DNA of *T.brucei* is dedicated to the approximately 1000 VSG genes present [Van der Ploeg *et al.*, 1982]. These genes are found in clusters throughout the genome in telomeric and non-telomeric positions. While the non-telomeric VSGs are never directly transcribed, all VSG genes in the 20 BSF telomeric sites have the potential to be expressed [Navarro *et al.*, 1996]. Even so, only a single telomeric expression site is transcribed at a give time producing trypanosomes that have a uniform VSG coat. The BSF VSG expression sites are 45-60 kb polycistronic units containing 10 genes termed expression site associated genes (ESAG). Some of the upstream ESAGs have been characterized, such as ESAG 6 and 7 which encode a membrane bound transferrin receptor [Ligtenber *et al.*, 1994] and ESAG 4 which encodes an adenylate cyclase molecule [Pays *et al.*, 1989]. Still other ESAG genes encode putative membrane associated molecules with unknown functions. The VSG gene is found at the 5' end of the expression site and is flanked by upstream by 70 bp imperfect repeats and downstream by telomeric repeats. Alteration in VSG expression occurs by activation of alternative expression sites or recombination of the VSG present in the actively transcribed site [Pays and Nolan, 1998]. The latter DNA recombination can be achieved via conversion of a segment or the entire VSG gene present, or via reciprocal recombination with one of the other VSG copies present in the genome.

1.7.2 Procyclins

During the transition from BSF into procyclic forms, the trypanosome sheds its VSG coat in favor of another unique surface membrane coat containing procyclin molecules. The major procyclic surface glycoprotein procyclin, gained its name from its stage specific expression and its proline rich content [Richardson *et al.*, 1988]. Several forms of procyclin exist in *T. brucei* and are termed either EP or GPEET depending on the type of repeat they possess. GPEET procyclins contain six Gly-Pro-Glu-Glu-Thr pentapeptide repeats and are phosphorylated at one or more of these threonine site [Butikofer *et al.*, 1997]. There are three forms of EP procyclin, EP 1, EP 2 and EP 3, which contains 29, 25 or 22 Glu-Pro dipeptide repeats respectively. All EP procyclins contain a putative GPI anchor site and both EP 1 and EP 3 contain N-glycosylation sites [Ruepp *et al.*, 1997]. The *T.brucei* spp. procyclin is expressed at 3-6 X 10^6 molecules per trypanosome and both EP and GPEET may be present on the surface concurrently. Different strains contain varying levels of EP and GPEET, and these proteins may be functionally distinct as they are not interchangeable as demonstrated by gene deletion experiments [Ruepp *et al.*, 1997]. EP forms of procyclin are predicted to have an extended rigid tertiary structure where the proline-glutamic acid repeats form a cylinder 0.9 nm in diameter and approximately 16 nm in length [Roditi *et al.*, 1989]. The glutamic acid-proline repeat structure found in PCF trypanosomes is resistant to almost all known proteases and thus likely forms a protective coat that insulates the trypanosome from the harsh environment of the tsetse fly midgut [Ferguson *et al.*, 1993]. It has also been suggested that this glycoprotein may be involved in localizing the binding of trypanosomes to different compartments within the tsetse fly or may function in parasite maturation [Roditi and Pearson, 1990; Stebeck *et al.*, 1994].

Several of the procyclin genes have been identified. *T.brucei* spp. contains tandem arrays of 2-3 procyclin genes copies at several locations throughout the genome [Mowatt, 1987; Koenig *et al.*, 1989]. These procyclin genes are transcribed by an alpha-amanitin resistant RNA polymerase [Koenig *et al.*, 1989] as an 8-10 kb polycistronic pre-mRNA transcript which contain procyclin associated genes (PAGs) and other elements between the

coding regions that are involved in posttranslational regulation. The stage specific mechanisms of procyclin expression which are still under investigation have been discussed elsewhere [Hehl and Roditi, 1994].

1.8 Minor Surface Molecules

While the stage specific coats of VSG and procyclin act in part to protect the parasite from the immune system and proteases, they have proven similarly obstructive to the identification and characterization of minor membrane molecules. With the popularization of the idea that functionally and immunogenically important molecules may be hidden beneath the surface coats made of procyclin or VSG, interest in the search for minor membrane-associated proteins has intensified, with fruitful results. To date more than twenty-five minor surface molecules have been identified, at least tentatively, including several surface receptors and transporters as well as a variety of invariant molecules.

1.8.1 Membrane Associated Receptors and Transporters

The flagellar pocket is a major site of localization for surface receptors, providing a heavily protected site at the point of flagellar emergence in which the plasma membrane invaginates [Langreth and Balber, 1975; Webster and Grab, 1988]. This site, which also lacks underlying microtubules, is believed to be the location of endocytotic activities such as nutrient uptake and secretion of various factors that entail membrane fusion processes. These processes would be hindered at other sites in the plasma membrane where underlying microtubules are present [Mclaughlin, 1987].

1.8.2 Transferrin Receptor

Transferrin is an essential trypanosome growth factor as it provides a source of iron [Webster and Grab, 1988]. This molecule is taken up from the bloodstream of the mammalian host through a high affinity transferrin receptor found in the flagellar pocket. The transferrin receptor is a heterodimeric complex that was initially partially purified by affinity chromatography on a human transferrin-coupled Sepharose column [Schell *et al.*, 1991]. The protein is encoded by the VSG expression site associated genes (ESAGs) 6 and 7, which together facilitate transferrin binding [Ligtenber *et al.*, 1994]. The component encoded by ESAG 6 is a 50-60 kDa protein modified by a phospholipase C sensitive GPI anchor at the C-terminus which attaches the heterodimer to the membrane. The ESAG 7 molecule is a 40 kDa protein that has an unmodified C-terminus [Steverding *et al.*, 1994]. Interestingly, ESAG 6 may be evolutionarily related to VSG having a high sequence similarity and similar N-terminal motifs. The conserved hypervariable motif is thought to be actively selected for protein variation, which could counteract the effects of anti-receptor antibodies, thus modifying possible complement mediated lysis [Zomerdijk *et al.*, 1991].

1.8.3 Adenylate Cyclase Receptors

ESAG 4 encodes one of several transmembrane adenylate cyclases located on the flagellum of *T. brucei*. Adenylate cyclases which facilitate the generation of the messenger cyclic AMP, have a receptor-like structure containing a large variable extracellular domain and a catalytic C-terminal domain similar to yeast adenylate cyclase [Pays *et al.*, 1989]. The different adenylate cyclases form a family of receptors that are predicted to bind specific ligands with their extracellular domain and may have distinct effects during parasite development. A calcium-activated adenylate cyclase has been described for *T. brucei* which is activated under stress conditions such as extreme cell density or an excessively acidic environment and may play a role in differentiation of BSF trypanosomes into procyclics [Rolin *et al.*, 1996].

1.8.4 Lipoprotein Receptor

Low density lipoproteins (LDL) and high density lipoproteins (HDL) are host derived factors specifically obtained at the flagellar pocket by receptor mediated endocytosis. Optimal growth of *T. brucei* is dependent on the LDL receptor as it provides access to cholesterol and other lipids carried within the LDL particle [Coppens *et al.*, 1995]. The LDL receptor is a 145 kDa protein that was first purified by affinity chromatography and is localized to the flagellar pocket [Coppens *et al.*, 1991]. While proteins binding with both high and low affinity are present, the LDL gene has not been found. The HDL particle is also taken up via receptor-mediated endocytosis. An unusually large and dense subclass of human serum HDL termed the trypanosome lytic factor (TLF) is responsible for lysis of susceptible strains [Hajduk *et al.*, 1994]. The trypanocidal activity of HDL has a rapid time course since 20-40 minutes after TLF addition, cells swell, lose mobility and lyse. This action occurs through lysosomal disruption and autodegradation after the TLF is taken up and delivered to the lysosome [Hajduk *et al.*, 1994]. Strains resistant to the lytic effects of TLF were found to have processing problems with the receptor binding TLF [Hager *et al.*, 1997].

1.8.5 Cytokine Receptors

Receptors for both tumor necrosis factor α (TNF α) and interferon γ (INF γ) have also been suggested by research on trypanosomes [Olsson *et al.*, 1993; Lucas *et al.*, 1994]. IFN γ is a host-derived cytokine that stimulates proliferation of trypanosomes, while a second host derived chemical, TNF α, has the opposite effect causing trypanosome lysis. Together these receptor ligand systems may act to regulate the growth of trypanosomes within the host such that a chronic infection can be established and is maintained. To this end, trypanosomes have been shown to secrete factors which direct this regulation by modulating both the TNF α and IFN γ cytokine level within the immune system of their host [Pays *et al.*, 1998].

1.8.6 Glucose Transporters

The presence of two distinct families of glucose transporters that are differentially regulated in *T. brucei* was first unearthed by the group of Baltz [Bringaud and Baltz, 1993]. They characterized a single locus that housed the glucose transporter genes that are preferentially expressed in either the BSF (termed tht 1) or the PCF (termed tht 2) trypanosomes. The locus contained a tandem array of six tht 1 genes preceding five tht 2 genes, and these distinct transporter gene families showed an 80 % sequence identity. The transporters have a similar structure containing 12 membrane spanning regions and a single N-terminal glycosylation site. Unlike their mammalian counterparts, the trypanosome glucose transporters contain an unusual cystine rich region in one of the hydrophobic loops and have the capacity to transport D-fructose [Bringaud and Baltz, 1992]. Striking distinctions between the two differentially regulated glucose receptor families appear in the extracellular looped domain that is thought to function in substrate binding. These differences may account for the much lower affinity of the receptor expressed in the BSF relative to the Tht 2 protein [Barrett *et al.*, 1995].

1.8.7 VSG and Procyclin Associated Minor Membrane Molecules

Several of the proteins that have genes located in the VSG or procyclin expression sites are in fact membrane-associated molecules that have expression patterns similar to their parent major surface glycoprotein. The ESAGs 1, 2 and 3, which were all identified by cloning, encode minor surface proteins of 360, 453 and 367 amino acids, that have uncharacterized functions [Alexandre et al., 1988]. While ESAG 1 is a glycosylated membrane associated protein, ESAG 2 is thought to be a GPI anchored surface protein localized to the cell surface [Pays *et al.*, 1989]. Due to the multiple ESAG 2 genes, deletion studies to determine the function of this protein have proven inconclusive [Carruthers *et al.*, 1996]. Further studies on ESAG 2 have turned up a related gene termed gene related to ESAG 2.1 (GrESAG 2.1) that is present and expressed from the PARP A locus during the PCF life

cycle stage [Berberof *et al.*, 1991]. The putative protein encoded by ESAG 3 is thought to possess a hydrophobic signal sequence and a GPI anchor, suggesting it interacts with the membrane [Alexandre *et al.*, 1988]. The 406 amino acid PAG 1 protein is hypothesized to associate with the plasma membrane via a GPI anchor and has significant similarities to the transferrin binding protein encoded by ESAG 6 and 7 [Koenig *et al.*, 1992]. As PAG 1 is preferentially expressed in the PCF it may be a transferrin receptor specifically adapted for the internal environment of the tsetse fly. Less is known about PAG 3 although it is thought to be a transmembrane protein that is exclusively restricted to the PCF trypanosome and has no known function.

1.8.8 Bloodstream Form Specific Invariant Surface Glycoproteins

Several invariant surface glycoproteins (Isg), including Isg 65, 75 and 100 are exposed to the surface and have been identified through biotinylation or radioiodination experiments. Isg 65 and Isg 75 genes are both present in multiple copies in the trypanosome genome while Isg 100 is encoded by a single gene [Ziegelbauer *et al.*, 1992; Ziegelbauer *et al.*, 1995; Nolan *et al.*, 1997]. The 65 and 75 kDa proteins of Isg 65 and Isg 75 share further similarities, having a large hydrophilic N-terminal extracellular region followed by a membrane spanning α-helix and finally a short cytoplasmic sequence [Ziegelbauer and Overath 1992]. These globular proteins are dispersed throughout the cell surface, hidden within the sea of VSG coating the trypanosome and are speculated to have a structural role due to their relative abundance of 5-7 X 10^4 molecules per cell [Ziegelbauer *et al.*, 1992]. The larger 100-kDa protein contains three domains as do the Isg 65 and Isg 75 molecules, but this is where the similarities end. The N-terminal domain containing three N-glycosylation sites is followed by a serine rich domain that contains 72 copies of a 17 amino acid repeat and a transmembrane spanning C-terminus that is thought to be transiently accessible to the external environment [Nolan *et al.*, 1997]. The function of Isg 100 is unknown; however, its cellular disposition largely within digestive vacuoles and

vesicles and to a lesser extent at the flagellar pocket, suggests it may be involved in the endocytotic processes.

Another Isg that might be involved in endocytosis at the flagellar pocket, is a glycoprotein termed Cb1-gp that was first identified from within a heterologous smear between 84 and 140 kDa by immunodetection [Brickman *et al.*, 1993]. Interestingly, it is transiently associated with the flagellar pocket before it is transferred to the Golgi apparatus and finally to the lysosomal compartment supporting this role for Cb1-gp. It is unclear if Cb1-gp and Isg 100 are related but like all the invariant surface glycoproteins discussed thus far; it is expressed exclusively in BSF trypanosomes. Studies on an additional invariant surface glycoprotein termed Isg 64 are still at a rudimentary stage, but its detection via radioiodination suggests it is surface exposed [Jackson *et al.*, 1993a].

1.8.9 Procyclic Form Specific Invariant Membrane Glycoproteins

A significant number of procyclic invariant membrane glycoproteins have been detected. These include the aforementioned Tht 2, PAG 1 and PAG 3 proteins as well as others such as procyclic stage surface antigen (Pssa)-2 [Jackson *et al.*, 1993b], Cram [Bastin *et al.*, 1994], gene related to ESAG (GrESAG) 2.1 [Berberof *et al.*, 1991] and kinetoplastid membrane protein (KMP)-11 [Stebeck *et al.*, 1995 and 1996; Bridge *et al.*, 1998]. Pssa 2 was identified by an immunopanning technique using an anti-PCF polyclonal antisera directed against proteins derived from procyclic cDNAs expressed in Cos cells. Pssa 2 is predicted to be a membrane spanning glycoprotein with an extended extracellular region and a short cytoplasmic domain with a repeat that is rich in proline. It is estimated that 3 X 10^5 copies of this 422 amino acid protein are present per trypanosome; however, the cellular location and function have still to be determined [Jackson *et al.*, 1993b]. Another protein termed Cram, which shares the cysteine rich repetitive motif of the LDL binding protein has also been identified [Bastin *et al.*, 1994]. Weather or not this flagellar pocket membrane protein is another form of the LDL receptor is uncertain at this point, as it is dissimilar in its expression pattern and it has a substantially smaller apparent molecular

weight than its potential partner. A gene related to ESAG 2, termed gene related to expression site associated gene 2.1 (GrESAG 2.1), is a protein encoded by a single gene and is preferentially expressed in PCF trypanosomes. The protein, first identified by cloning of its gene, is thought to be associated with the cell surface membrane by a GPI anchor [Berberof *et al.*, 1991]. Like all the invariant surface glycoproteins discussed, the function of this protein is unknown. Perhaps the most abundant of the minor membrane associated molecules expressed predominantly in the PCF trypanosomes is the kinetoplastid membrane protein -11 (KMP-11). With up to 1 X 10^6 copies per trypanosome, KMP-11 rivals the cellular abundance of procyclin (3 - 6 X 10^6 per cell), but it is restricted in location mainly to the flagellum and to distinct areas at the flagellar base [Stebeck *et al.*, 1996]. This 11-kDa protein may perform an essential role in the procyclic form trypanosome, as it certainly shows elevated expression in this life cycle stage [Stebeck, 1996]. The trypanosome KMP-11 genes have recently been cloned and the locus sequenced, showing a tandem array of four genes [Bridge *et al.*, 1998]. These experiments as well as KMP-11 gene deletion studies will be discussed further as the KMP-11 genes are the focus of this thesis.

CHAPTER 1: Identification and Characterization of the Kinetoplastid Membrane Protein-11 Gene Locus of *Trypanosoma brucei*

INTRODUCTION

The kinetoplastid membrane protein-11 (KMP-11) was first identified in *Leishmania donovani* [Jardim *et al.*, 1991] and subsequently in a variety of kinetoplastid parasites including African trypanosomes, American trypanosomes, Crithidia, Leptomonas and Phytomonas [Stebeck *et al.*, 1995]. The ubiquitous distribution of the 11-kDa protein in kinetoplastids was initially elucidated using an immunologically cross reactive monoclonal antibody (mAb) specific for what was originally termed the leishmania lipophosphoglycan-associated protein [Tolson *et al.*, 1989]. The protein was not found in mammalian cells or in other non-kinetoplastid protozoan parasites tested. Immunoblot analysis of the expression of the KMP-11 protein throughout the African trypanosome life cycle revealed that it was present in all life cycle stages in varying levels. Much higher expression was detectable in the insect form PCF trypanosomes than in the mammalian-infective BSF trypanosomes [Stebeck *et al.*, 1995]. These characteristics were further supported by studies in which *T.congolense* BSF were transformed to PCF and a dramatic increase of KMP-11 expression was measured four hours after initiating transformation. The expression plateaued 48 to 72 hours later. Dot blotting analysis showed that the KMP-11 protein extracted mainly in the detergent phase of Triton X-114. Since the 11-kDa protein was found associated with the membrane (non-aqueous phase), its surface accessibility on the trypanosome was examined by biotin labeling experiments using a enhanced sensitivity chemiluminescence procedure [Stebeck *et al.*, 1995]. No KMP-11 protein was detected in several experiments using two distinct biotinylation reagent systems. Immunofluorescence on living *T.congolense* and *T.b.rhodesiense* was also negative supporting the notion that the protein was not surface accessible. These results from African trypanosomes mirror those observed in Leishmania where both surface biotinylation studies and immunofluorescence on live organisms using KMP-11 specific mAbs were negative.

Where immunofluorescence on live trypanosomes failed, experiments with acetone-permeabilized cells produced some striking results [Stebeck *et al.*, 1995]. Monoclonal antibodies against KMP-11 produced a distinctive string-like pattern of fluorescence along the flagellum, ending with two bright spots at the reservoir region of the flagellar base (Figure 4).

Once the protein had been purified via reversed-phase chromatography using an octyl-Sepharose column, further characterizaion of the molecule was undertaken. No carbohydrates were detected on the purified protein in several enhanced chemiluminescence biotin hydrazide procedures. Gravimetric analysis indicated that KMP-11 was expressed at approximately 2 X 10^5 to 1 X 10^6 molecules per trypanosome, making it the third most abundant trypanosome membrane molecule characterized to date, only surpassed by the stage specific proteins procyclin and VSG [Stebeck *et al.*, 1995]. Endopeptidase Lys-C digestion followed by amino acid microsequencing was performed on the trypanosome KMP-11 protein revealing an internal sequence that was 86% identical to a section of sequence in the *L.donovani* KMP-11 molecule. Due to this high degree of similarity, the *L.donovani* KMP-11 DNA sequence was used as a heterologous probe in screening of a *T.brucei* genomic DNA cosmid library [Stebeck *et al.*, 1995]. A positive cosmid clone containing two-thirds of the KMP-11 gene sequence, including the 3' end, was isolated and RT-PCR was used to obtain the missing gene sequence. The trypanosome KMP-11 protein was found to have an encoding gene of 279 bp, producing a protein of 92 amino acids with a calculated molecular mass of 11, 078 Da, a pI of 6.0 and a net charge of -2 at physiological pH. Alignment of the *T.b.rhodesience* and *L.donovani* KMP-11 protein sequences showed an 82% sequence identity [Stebeck *et al.*, 1995]. One of the most highly conserved regions between the two proteins was also the location of the epitope recognized by the mAbs specific for KMP-11 (L98/L157). This serendipidous occurence made the initial cross species identification of KMP-11 possible.

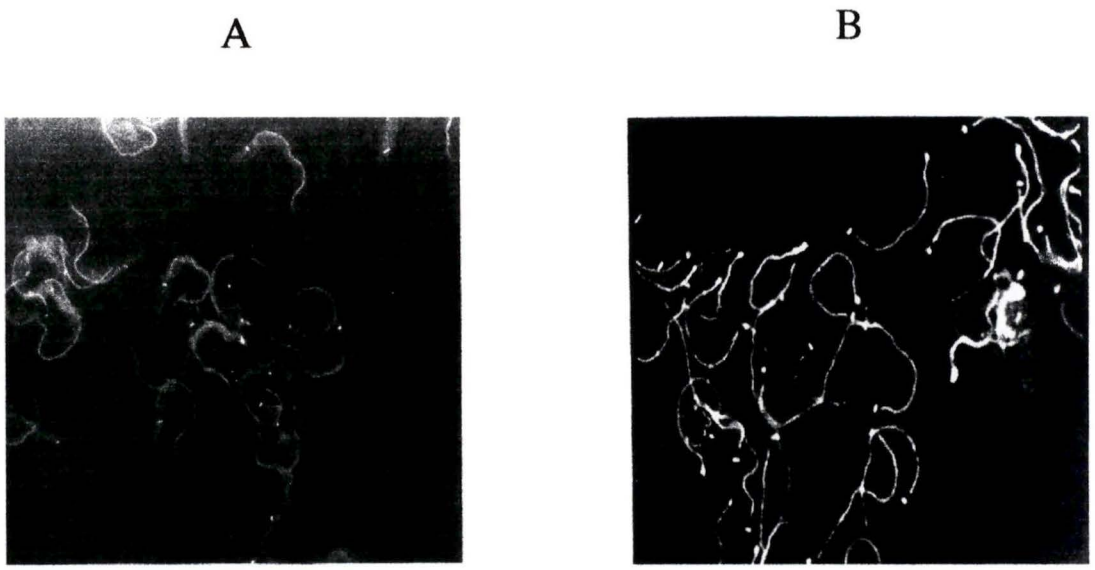


Figure 4: Immunofluorescence analysis of KMP-11 expression in *T.congolense* IL-3000 PCF. (A) standard fluorescence microscopy, (B) confocal laser scanning immunofluorescence microscopy. The anti-KMP-11 mAb L157 was used as the primary antibody in both experiments.

Both Leishmania and trypanosome KMP-11 proteins are predicted to assume a helix-turn-helix motif in their secondary structure [Jardim *et al.*, 1995; Stebeck *et al.*, 1996]. The two helices of 31 residues at the amino-terminus and 50 residues near the carboxy-terminus were predicted to be separated by a relatively large random coil, with a smaller random coil present at the carboxy-terminus (Figure 5). Each of the alpha-helices contains a region of hydrophobic residues (shaded), which creates an area of high hydrophobicity, while the rest of the protein is comparatively hydrophilic in nature. This amphipathic molecule could interact with the lipid bilayer through these hydrophobic regions, leaving the hydrophilic areas exposed to the surroundings. This type of association with the phospholipid bilayer underscores how KMP-11 could associate with the membrane without requiring the use of a lipid anchor or hydrophobic polypeptide to hold it in place. The anti-KMP-11 mAb epitope lies across a large part of the central random coil in a relatively accessible portion of the protein, facing away from the lipid bilayer.

Southern blot analysis produced evidence supporting the existence of only a single KMP-11 gene copy in the trypanosome genome [Stebeck *et al.*, 1996]. This indicated that trypanosome KMP-11 might be an ideal candidate for gene deletion mutagenesis experiments. As a prelude to attempting knockout mutagenesis for studies of KMP-11 function, it was necessary to analyze the KMP-11 coding and flanking regions from *T.brucei* genomic DNA in more detail. In this chapter, I show that the KMP-11 locus contains four closely related KMP-11 gene repeats in a tandem array. The data suggests the presence of regulatory sequences and gives insight into the duplication events involved in the production of this locus [Bridge *et al.*, 1998].

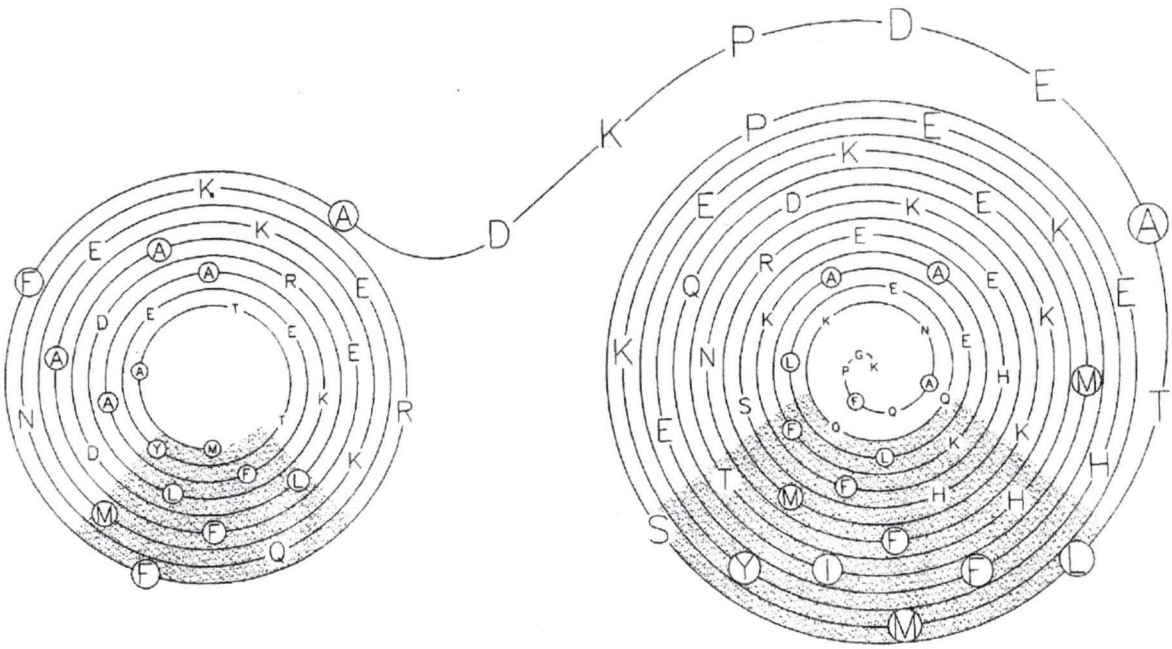


Figure 5: A schematic representation of the secondary structure of KMP-11 [Stebeck *et al,* 1996] predicted using the algorithn of Garnier [Garnier *et al*, 1974]. The circled residues indicate hydrophobic amino acids, while the shaded region indicates the hydrophobic face of the KMP-11 helices.

MATERIALS AND METHODS

Parasites. Procyclic culture forms (PCF) of *T.b.brucei* 348 TI, *T.b.brucei* 348 T2, and *T.b.brucei* 427.01; epimastigotes of *T.cruzi* STIB 68 MA, *T.cruzi* STIB 745, *T.cruzi* 1457; and promastigotes *L.major* YS119 (provided by Dr. Isabel Roditi, Institut fur Allgemeine Mikrobiologie, Universitat Bern, Bern, Switzerland) were maintained in culture at 27°C in a 5% CO_2 in air atmosphere. Parasites were grown in minimal essential medium (MEM) containing 25 mM HEPES, 0.2 mM hypotanthine, 2.5 mg/ml hemin, 60 mM proline, 2 mM glutamine, 1% non-essential amino acids, 50 µg/ml gentamycin, and 10% fetal bovine serum (FBS). For healthy log-phase cultures, this procyclic culture medium was changed either every other day or whenever the medium turned yellow, indicating increased acidity.

Isolation of trypanosome genomic DNA . Wild type trypanosomes were grown to approximately $3x10^7$ per ml and harvested by centrifugation at 2000x gfor 5 minutes at room temperature on a Beckman TJ-6R tabletop centrifuge (Beckman, Palo Alto, CA, USA). The cells were resuspended in an Eppendorf tube in 500 µl of TELT (50 mM Tris-HCl pH 8.6, 2.5 mM EDTA, 2.5 M LiCl, l4% Triton X-100) by inverting the tube gently several times and then incubating for 5 minutes at room temperature. Five hundred microliters of pre-equilibrated phenol/chloroform (Anachemia, Montreal, QE, Canada) was added and the suspension was mixed gently on a Lab Quake™ (Fisher Scientific, Nepean, ON, Canada). The mixture was then centrifuged for 5 minutes at 15,000x gat room temperature in a microcentrifuge and the upper phase was transferred into a new Eppendorf tube. To precipitate the genomic DNA, 1.0 ml of ice-cold 95% ethanol was added to the supernatant and the tube was centrifuged at 15,000 x g for 15 minutes at 4°C. The supernatant was discarded, the pellet was washed with 70% ethanol and then air dried. Once the pellet had been dissolved in 100 µl of TER (50 mM Tris pH 8, 62.5 mM EDTA, 20 µg/ml RNase A) and incubated for 30 minutes at 37°C, the DNA concentration and

purity were determined by measuring the OD at 260 nm and 280 nm. The conversion was 1.0 OD_{260} = 50 ng/ul. The isolated genomic DNA was stored at 4°C.

Polymerase chain reaction (PCR). The nucleotide sequence of the gene encoding the kinetoplastid membrane protein-11 (KMP-11) [Stebeck *et al.*, 1996] was used to design forward and reverse PCR primers for the amplification of the entire KMP- 11 coding region from trypanosome genomic DNA. All PCR primers were designed using Oligo 4.0 software (National Bioscience, Inc., Plymouth, MN, USA) to eliminate possible primer dimer and hairpin loop formations. The PCR amplification mixture (100 µl) contained 25 pmol of each of the forward and reverse primers (Gibco BRL, Burlington, ON, Canada), 2 mM each of dATP, dCTP, dGTP and dTTP, 1x Taq DNA polymerase buffer, 2.5 units of Taq DNA polymerase (all from Pharmacia, Piscataway, NJ, USA) and 100 ng trypanosome genomic DNA. The forward primer #4 [5'-ATG GCC ACC ACA TAC GAA G -3'] contains the ATG start codon and was complementary to the 5' end of the KMP-11 coding region. The reverse primer #5 [5'-TCA TTT TCC GGG GAA CTG -3'] contains a stop codon, complementary to the 3' end of the KMP-11 coding region. The PCR conditions were : 95°C for 5 minutes (denaturing), followed by 30 cycles of the following: 95°C for 1 minute (denaturing), 58°C for 30 seconds (annealing), 74°C for 1 minute (extension). Finally an extension was continued at 74°C for 5 minutes prior to holding at 4°C. The amplified KMP-11 gene product was isolated and used in library screening and Southern blot analysis as a probe.

Bacteriophage P1 library screening[1]. A P1 bacteriophage high density filter library (SM7 No 8), containing genomic DNA of *T.brucei* strain TREU 927/4 [Turner, 1990] was obtained from Drs. Sara Melville and Vanessa Leech (Laboratory for Parasite Genome

[1] Library screening was performed by Caroline Stebeck, Department of Biochemistry and Microbiology, University of Victoria, Victoria BC, Canada.

Analysis, Department of Pathology, Cambridge University, Cambridge, England). The *T.brucei* genomic DNA library was constructed using the bacteriophage P1 cloning vector, pAD10SacBII, as previously described [Pierce *et al.*, 1992]. The genomic DNA fragments of 60 to 70 kb were generated by partial digestion of the trypanosome genome with *Sau*3A*I* and these fragments were subsequently cloned into the *Bam*HI site of the P1 vector. The library was grown in twenty three 96-well microtiter plates and the colonies from each well were transferred directly to a high density filter membrane (Hybond N, Amersham, England). The high density filter was probed with an alpha ^{32}P-labeled trypanosome KMP-11 genomic PCR gene amplification product. Prehybridization and hybridization with the library, as well as the filter washing and autoradiography were performed as described for Southern blot analysis (see below).

Processing of bacteriophage P1 clones. The P1 clones, 3D12 and M1A7, that hybridized most strongly with the KMP-11 probe were selected for further analysis. DNA was isolated from these clones using conditions described by the Laboratory for Parasite Genome Analysis. In brief, *E.coli* cells containing the cosmids were regrown from an agar slant on LB plates containing ampicillin (25ug/ml) and a single colony was then used to innoculate a liquid LB culture. A new culture was innoculated using 1/100 volume of the overnight culture and was incubated for 30 minutes at 37°C shaking at 250 rpm. The P1 lytic operon was induced by adding IPTG to a final concentration of 1mM and the incubation was continued for another five hours. The cells in culture were harvested by centrifugation, resuspended in gET buffer (50mM glucose, 10mM EDTA pH8, 25mM Tris HCl pH8) containing 10 mg/ml RNAse and subjected to an alkaline lysis cosmid preparation procedure [Birnboim and Doly, 1979]. After a phenol/chloroform extraction the bacteriophage vector was ethanol precipitated and washed before being resuspended in TE buffer. The DNA was denatured (95°C, 3 minutes) and dotted onto Zeta-Probe® membrane (BioRad, Hercules, CA, USA) according to the manufacturer's instructions. These were screened, using the homologous KMP-11 probe under conditions outlined for

the Southern blot analysis, to verify positivity. One positive trypanosome P1 clone (3D12) was selected for further analysis. The cosmid insert size was approximately 60 to 80 Kb. To obtain a more manageable insert size the DNA was isolated from this clone, digested with a variety of enzymes (see Southern blot analysis) and separated on a 0.6% agarose gel (Promega, Madison, WI, USA). The DNA was transferred to Zeta-Probe® membrane (BioRad, Hercules, CA, USA) and probed with the trypanosome KMP-11 probe as described for the Southern blot analysis (see below).

Agarose gel electrophoresis. Unless otherwise indicated, 1% agarose gels were used. Molecular biology grade agarose LE (Promega, Madison, WI USA) was dissolved in Tris acetate (0.04 M Tris acetate, 0.001 M EDTA pH 8.5) buffer to prepare the gel. *Hind* III digested Lambda DNA standards (New England Biolabs, Beverly, MA, USA) or a 1 kb DNA ladder (Gibco BRL, Burlington, ON, Canada) served as size standards for each gel. The DNA was visualized and photographed using a Fotodyne transilluminator and camera (BioCan Scientific, Mississauga, ON, Canada) after being stained with 1.25 μM ethidium bromide (Sigma, St. Louis, MO, USA).

Southern blot analysis. For Southern blot analysis, 8 μg of DNA from the original P1 bacteriophage clone was digested overnight at 37°C with 100 units of each of *Hind* III, *sph* I, *Bam* HI, *sal* I, *Nde* I and *Nco* I (New England Biolabs, Beverly, MA, USA), followed by a further 4 hour incubation at 37°C with an additional 100 units of enzyme to ensure complete digestion. The digests were electrophoresed for 12 hours at 15 V on a 0.6% agarose gel, along with a *Hind* III digested lambda DNA-size standard (New England Biolabs, Beverly, MA, USA). The DNA restriction digests were transferred by vacuum blotting (model 785, BioRad, Hercules, CA, USA) on to Zeta-Probe® membrane (BioRad, Hercules, CA, USA) with slight modifications to the manufacturer's instructions. Briefly, the agarose gel for Southern blot analysis was treated with 0.25N HCl (1 x 15

minutes), 1.5 M NaCl/0.5 M NaOH (2 x 15 minutes) and 1.5 M NaCl/0.5 M Tris-HCl pH 7.5 (1 x 30 minutes) prior to transfer in 0.5 M NaOH/0.6 M NaCl. Approximately 25 ng of the genomic PCR fragment corresponding to the entire KMP-11 coding region (see genomic PCR methods) were labeled with [α-^{32}P] dATP (specific activity 3000 Ci/mmol) using the Multiprime DNA labeling system (Amersham, Oakville,ON, Canada) according to the manufacturer's instructions. The Southern blots were treated with prehybridizion buffer for 5 minutes at 50°C in 0.25 M Na_2HPO_4 pH 7.2/7% SDS and then were incubated overnight at 50°C with mild agitation in the same buffer containing the labeled probe (1.0×10^{-6} cpm/ ml; specific radioactivity 2.0×10^{-9} cpm/μg). The membranes were washed 2 times for 1 hour each at 50°C with 200 ml of 20 mM Na_2HPO_4 pH 7.2/5% SDS, followed by 2 washes for 20 minutes each at 50°C with 200 ml of 20 mM Na_2HPO_4 pH 7.2/1% SDS. The membranes were air-dried and subsequently autoradiographed using a STORM 820 storage phosphor imaging system (Molecular Dynamics, Sunnyvale, CA, USA).

Cloning and transformation. *E.coli* strain DH5α (Gibco BRL, Burlington, ON, Canada) and the tetracycline resistant strain XL-1 Blue (Stratagene, La Jolla, CA, USA) were used in conventional [Hanahan, 1983] and high voltage [Dower *et al.*, 1988] electroporetic transformation experiments. Standard gene cloning techniques were used [Sambrook *et al.*, 1989]. Bacterial cultures were grown in either Luria-Beitani (LB) medium (1% bacto-tryptone, 0.5% bacto-yeast extract, 0.17 M NaCl, pH 7.0) or 2 x YT medium (1.6% bacto-tryptone, 1% bacto-yeast extract, 86 mM NaCl, pH 7.0) [Sambrook *et al.*, 1989]. The insert DNA fragments were recovered from low-melting temperature agarose using Qiaex II gel extraction Kit (Qiagen, Chatsworth, CA, USA) and cloned into either the Bluescript™ SK+ vector (Stratagene, La Jolla, CA, USA) or the M13mp19 vector (Pharmacia Biotech, Baie d'Urfe, QE).

RNA isolation. RNA was isolated from 3 x 10^8 procyclic culture form trypanosomes or leishmania promastigotes grown to log phase at approximately 1 - 5 x 10^6/ml. The parasites were harvested by centrifugation using a Beckman TJ-6R benchtop centrifuge (Beckman, Palo Alto, CA, USA) at 2000x gfor 5 minutes at room temperature, the supernatants were decanted, and the cells were resuspended in 200 µl of the residual liquid. The RNA extraction mixture was prepared as follows: 50 µl of 10% SDS (100 g SDS, in 900 ml of sterile DEPC treated water, disolved by stirring for 1 hour at 68°C, adjusted to pH 7.2 and the volume adjusted to 1 L), 450 µl of NTE (1 M NaCl, 50 mM EDTA, 100 mM Tris pH 7.5) and 500 µl of phenol were heated together to 80-90°C in a 1.5 ml Eppendorf tube until the solution became monophasic. The resuspended parasites were pipetted into the hot phenol RNA extraction mixture and gently mixed for 3 minutes by inverting the tubes. After the mixture had been centrifuged in a conventional microcentrifuge at 12000 g, for 5 minutes at room temperature, the upper aqueous phase was transferred to a new Eppendorf tube and extracted once with 500 µl of phenol and then once with 500 µl of chloroform. The resulting upper aqueous phase was transferred to a sterile 2 ml Eppendorf tube and 1400 µl of 95% ethanol were added. The RNA was precipitated by centrifugation at 12000 g, for 30 minutes at 4°C and then the pellet was washed with 500 µl of 70% ethanol under the same centrifugation conditions. The RNA pellet was air dried and disolved in 100 µl of sterile distilled water . The OD_{260} was measured to determine the RNA concentration. The conversion used was 1 OD_{260} = 40 ng/ul. The RNA was stored at -20°C (short term) or -70°C (long term).

Northern blot analysis. To analyse the KMP-11 mRNA transcripts of several kinetoplastid parasites, northern blot analysis was performed. A 1.4% agarose gel in 10 mM NaHPO4 buffer was poured. A denaturing solution was prepared which contained the following: 105 µl of DMSO, 30 µl deionized glyoxal (30%) and 4.2 µl $NaHO_4$ (500 mM, pH 6.85). After two volumes of this denaturation buffer were added to 4 µg of RNA dissolved in 1

volume of water, the mixture was incubated for 8 minutes at 50°C and then transferred to ice. The loading buffer (50% glycerol, 5 x phosphate buffer, 0.4% bromophenol blue) was then added at 0.39 volumes of the loading buffer per volume of the denatured RNA. The agarose gel was electrophoresed for 2 to 3 hours at 4 V/cm. To keep a steady pH, the buffer was circulated by stirring with a stir bar in each end of the gel apparatus during running. After running the gel was photographed on a UV light box and capillary blotted to transfer the RNA to a membrane. For the capillary blot, a tray was filled with 20 x SSC and a glass plate covered with a piece of Whatman 3M filter paper that dipped into the blotting buffer was used as a platform. A pre-wetted piece of Whatman 3M filter paper slightly larger than the gel was placed onto the first layer and the gel was placed on top of it so that no bubbles were between the paper and the gel. A pre-wetted sheet of Zeta-Probe® membrane (BioRad, Hercules, CA, USA), cut to the dimensions of the gel was then placed (avoiding bubbles) on the gel. Three pieces of filter paper cut to the size of the gel, a stack of paper towels 5 cm in height, and glass plates were then piled on top of the membrane and the transfer proceeded for 12 to 16 hours. After blotting, the apparatus was carefully dismantled and the membrane was washed briefly in 2 x SCC buffer before being air dried. The RNA was then fixed on to the membrane by baking the membrane at 80°C for 30 minutes. The prehybridization, hybridization, and washing steps were performed as described for the Southern blot procedure. Autoradiography was performed using a STORM 820 storage phosphor imaging system (Molecular Dynamics, Sunnyvale, CA, USA).

Library generation for DNA sequencing. A comprehensive library for automated DNA sequencing was produced from the 3500 bp P1 phage clone insert containing the KMP-11 locus using a method slightly modified from Povinelli [Povinelli *et al.*,1993]. Large

amounts of the insert were generated from PCR reactions using the universal forward and reverse primers which flank the Bluescript™ multiple cloning site. Amplified PCR fragments were recovered from low melting temperature agarose gels using Wizard™ PCR Preps purification system (Promega, Madison, WI, USA). Fifty micrograms of the PCR product in nebulization buffer (0.1 M Tris pH 8, 0.015 M MgCl, 10% glycerol in a total volume of 2.0 ml) were randomly disrupted using a medical nebulizer [Bodenteich *et al.*,, 1994] at 30 psi for 2 minutes. Nebulized fragments were incubated with 100 units of mung bean nuclease to remove the single stranded internal gaps and then repaired using 40 units of T4 DNA polymerase and 25 units of Klenow fragment of DNA polymerase (all from New England Biolabs, Beverly, MA, USA). DNA fragments ranging from 0.75 to 1.0 kb were separated on a 0.7% Nusieve agarose gel and purified using the Qiaex II gel extraction kit (Qiagen, Chatsworth, CA, USA). These inserts were then ligated into the M13mp19 vector (Pharmacia Biotech, Baie d'Urfe, QE) which had been previously digested with *Hinc* II. The ligated DNA was electroporated into *E.coli* DH5α cells using a BioRad gene pulser at resistance - 200 Ω; voltage 18 kV; capacitance - 25 μF (BioRad, Hercules, CA, USA). The electroporated cells were plated in top agar onto 2 x YT plates containing 5-Bromo-4-chloro-3-iodolyl β-D-galactopyranoside (X-gal) and Isopropyl β-D-thiogalactopyranoside (IPTG) (both from Sigma, Oakville, ON, Canada) and incubated overnight. The clear plaques were picked and directly PCR tested using M13 universal forward and reverse primers to ensure that they contained inserts. The vectors containing inserts were subsequently grown with host *E.coli* in liquid LB culture containing 25μg/ml kanamycin (Sigma, Oakville, ON, Canada) overnight at 37°C with vigorous shaking. The cultures were centrifuged and the supernatants were extracted using 20% PEG/2.5M NaCl and then phenol-chloroform. The templates were then ethanol/sodium acetate precipitated and resuspended in TE in preparation for DNA sequencing.

DNA Sequencing. Automated cycle sequencing was performed following the instructions in the ABI PRISM Cycle Sequencing Ready Reaction Kit With AmpliTaq DNA

polymerase, FS (Perkin Elmer, Mississauga, ON, Canada). Briefly, the templates were diluted to 100 ng/µl and approximately 1 µl of DNA was added to the A and C tubes and 2 µl to the G and T tubes. The dye primer PRISM (ABI) mix was added to the four PCR tubes as follows: 4 µl A, 4 µl C, 8 µl G and 8 µl T respectively. The reactions were performed using the following PCR program: (1) 94°C for 2 minutes; (2) 15 cycles of 95°C for 30 seconds, 55°C for 1 minute and then 70°C for 1 minute; (3) 15 cycles of 95°C for 30 seconds, 70°C for 1 minute, a rapid ramp to 95°C; (4) 20°C for 10 minutes and finally a hold at 4°C. The amplified DNA was preciptated using sodium acetate/ethanol, dried using a Speedvac™ (Savant, Hicksville, NY, USA) and dissolved in 4.1 µl of formamide. The formamide DNA solution was denatured at 90°C and loaded into a 5% acrylamide/urea gel (25g urea, 21.25ml water, 6.25ml of 40% acrylamide, 5ml of 45mM Tris-borate with 1mM EDTA) and sequenced on the ABI 373 sequencer (Perkin Elmer, Mississauga, ON, Canada). Seventy two individual positive clones from the KMP-11 locus M13 bacteriophage library were sequenced.

Comprehensive DNA sequence assembly. The sequence data was aligned and assembled using Seqman™ sequence alignment software (DNASTAR, Madison, WI, USA). The window size was 12 bp with stringency set at 70%. Seventy-two individual sequencing reads averaging 550 bp per read were assembled and gave an 11.5 fold sequence redundancy. This allowed accurate assembly of the highly repetitive KMP-11 sequence.

Dot matrix sequence analysis. The known cDNA sequences of KMP-11 taken from NCBI and Genbank databases were used to search for genes and pseudogenes. Dot matrix comparisons were performed using the Geneworks™ dot matrix analysis program (IntelliGenetics Inc., Mountain View CA) where a dot corresponds to a window size of 20 bp with a minimum number of 14 matches within that window and an offset of 10 bp. Characterization of the promoter region was performed using both the Gene Features

Search and Multiple Sequence Alignments provided by the Human Genome Center, and Baylor College of Medicine respectively. Open reading frames of greater than 200 bp were found using Seqman™ sequence alignment software (DNASTAR, Madison, WI, USA).

RESULTS

P1 bacteriophage library screening. The *T.brucei* strain TREU 927/4 P1 library containing genomic DNA was screened using the 279 bp trypanosome genomic KMP-11 probe to obtain the KMP-11 gene and DNA flanking regions. The screening identified two *T.brucei* clones, 3D12 and 1A7, that hybridized to the KMP-11 probe. Both these clones were further characterized by dot blot and Southern blot analysis. An autoradiograph of the high density library filter containing the positive 3D12 clone is shown (Figure 6).

Dot blot analysis of the positive clones. To confirm the positive results obtained from the library screening, both P1 bacteriophage clones (3D12 and 1A7) were re-tested by dot blot analysis. After the bacteriophage clones were grown, the insert containing pAd10SacBII vectors were isolated and the DNA was quantified. Aliquots of the P1 clone DNA and control DNA were serially diluted and dotted onto a Zetaprobe™ membrane (BioRad, Hercules, Ca, USA) before hybridization with the ^{32}P-labelled 279 bp trypanosome genomic KMP-11 probe. The 3D12 and 1A7 clone DNA hybridized with similar strength to the positive control DNA (PCR amplified KMP-11 gene fragment) within the range of DNA bound to the membrane for the experiment (Figure 7, lanes 1-3). There was strong hybridiztion with 500 ng to 50 ng of coated DNA while at quantities of 5 ng and lower the signal was weak. Conversely the signal elicited from the negative control (PCR amplified glycerol-3-phosphate dehydrogenase DNA fragment) was not detected even at 500 ng of coated DNA (Figure 7, lane 4). These results indicated that the inserts carried by the P1 bacteriophage clones 3D12 and 1A7 likely contained at least a portion of the KMP-11 gene.

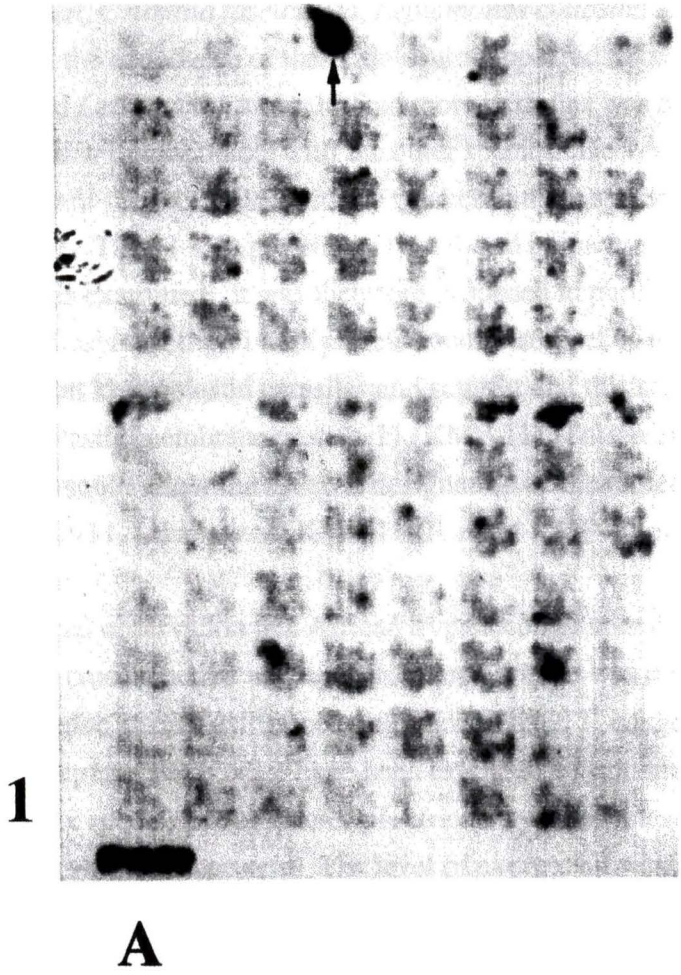


Figure 6: Autoradiograph of the trypanosome P1 bacteriophage library high density filter screened with the ^{32}P-labeled 279 bp KMP-11 gene from *T. b.rhodesiense* . The arrow indicates the hybridizing trypanosome clone 3D12 that was selected for further analysis. To identify clones the filter was oriented such that the letters A to H extend along the bottom and numbers 1 to 12 run vertically. The clone 3D12, indicated by the arrow is located in row 12, column D and is the third spot in the D12 block.

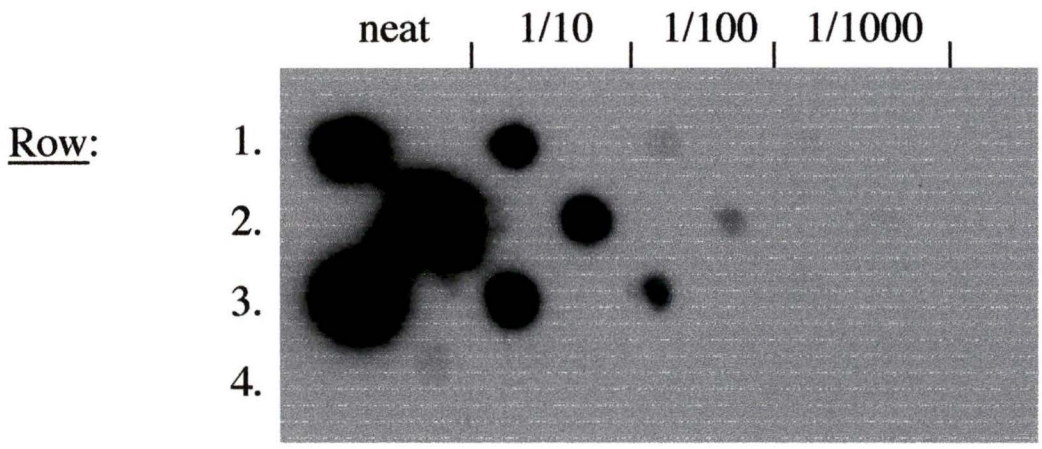


Figure 7. Dot blot analysis of the KMP-11 positive clones selected from screening the P1 bacteriophage library containing *T.b.brucei* strain 927 genomic DNA. The isolated P1 clone and control DNA were serially diluted from a standard solution of 100 ng/ul and aliquots of 5 ul were used.

Row 1- DNA from P1 clone 1A7
Row 2 - DNA from P1 clone 3D12
Row 3 - DNA from PCR-amplified KMP-11
Row 4 - DNA from PCR-amplified glycerol-3-phosphate dehydrogenase gene. [Stebeck *et al.*, 1996]

Isolation of a genomic DNA fragment containing the KMP-11 gene and flanking sequences. The P1 bacteriophage 3D12 clone was selected for further analysis due to the strong hybridization signal that it gave during the library screen and subsequent screenings. To generate a KMP-11 genomic DNA fragment of a more managable cloning size, the 65 kb insert was restriction digested and analysed by Southern blotting. The 3D12 clone DNA was digested using the endonucleases *Hind* III, *Sph* I, *Bam HI*, *Sal* I, *Nde* I and *Nco* I (New England Biolabs, Beverly, MA, USA), which do not have restriction sites within the KMP-11 gene. The digests were electrophoresed, transferred to Zetaprobe™ membrane (BioRad, Hercules, Ca, USA) and probed using the 279 bp trypanosome KMP-11 probe. The Southern blot showed hybridizing bands with sizes of greater than 10,000 bp for the *Hind*III, *Sph*I, *Nde*I and *Nco*I digests (Figure 8, lanes 1, 2, 5 and 6) while the *Bam*HI and *Sal*I digests showed more promising bands at approximately 3500 bp and 9500 bp respectively (Figure 8, lanes 3 and 4). Bands equivalent to the 3500 bp *BamHI* and 9500 *Sal*I bands seen on the Southern blot were subsequently isolated from an agarose gel and subcloned into the pBluescript™ cloning vector (Stratagene, La jolla, Ca, USA). The resulting vector constructs were termed pBS KMP-11 3c (BamHI fragment) and pBS KMP-11 2 (SalI fragment).

Characterization of the KMP-11 constructs. A PCR amplification of the KMP-11 gene was performed on the newly cloned constructs and the original P1 clone to verify the presence of the correct insert. KMP-11 forward primer #1 and the KMP-11 reverse primer #2 (Figure 18, Page 55; and Table 1, Page 56) were used to amplify the KMP-11 genes that were resident in the *sal*I, *Bam*HI and P1 inserts (Figure 9). The P1 bacteriophage clone 3D12 showed a single PCR product at approximately 279 bp which corresponded to the size of the KMP-11 coding region (Figure 9, lane 3). The reactions containing the pBS KMP-11 2 and pBS KMP-11 3c constructs as templates showed a similar sized band (279 bp) but also showed a series of bands at 850 bp, 1400 bp and 1950 bp (Figure 9, lane 1 & 2) which were also observed in subsequent experiments. The 279 bp KMP-11 gene

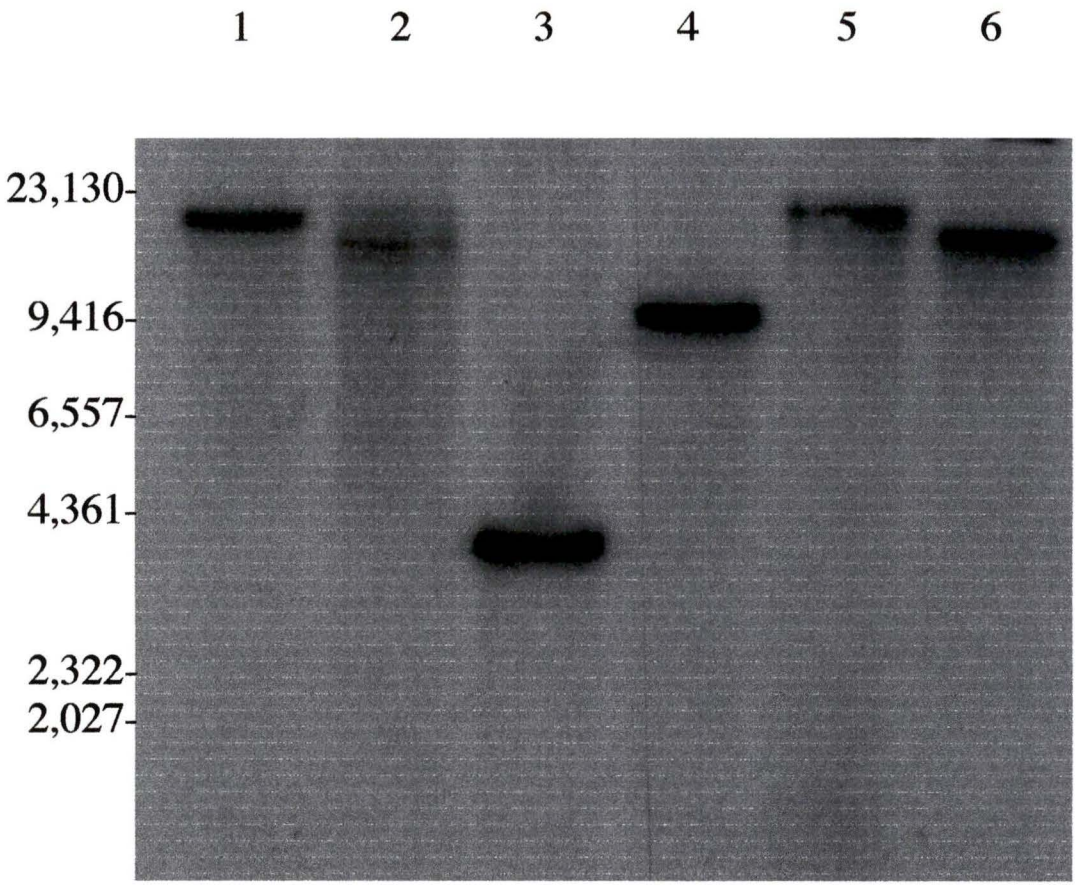


Figure 8: Southern blot analysis of the P1 bacteriophage clone 3D12 containing the KMP-11 gene locus, probed with the ^{32}P-labeled PCR amplified KMP-11 gene product. Lanes 1 to 6 represent bacteriophage P1 clone 3D12 DNA digested with the restriction endonuclease *Hind*III, *Sph*I, *Bam*HI, *Sal*I, *Nde*I or *Nco*I respectively.

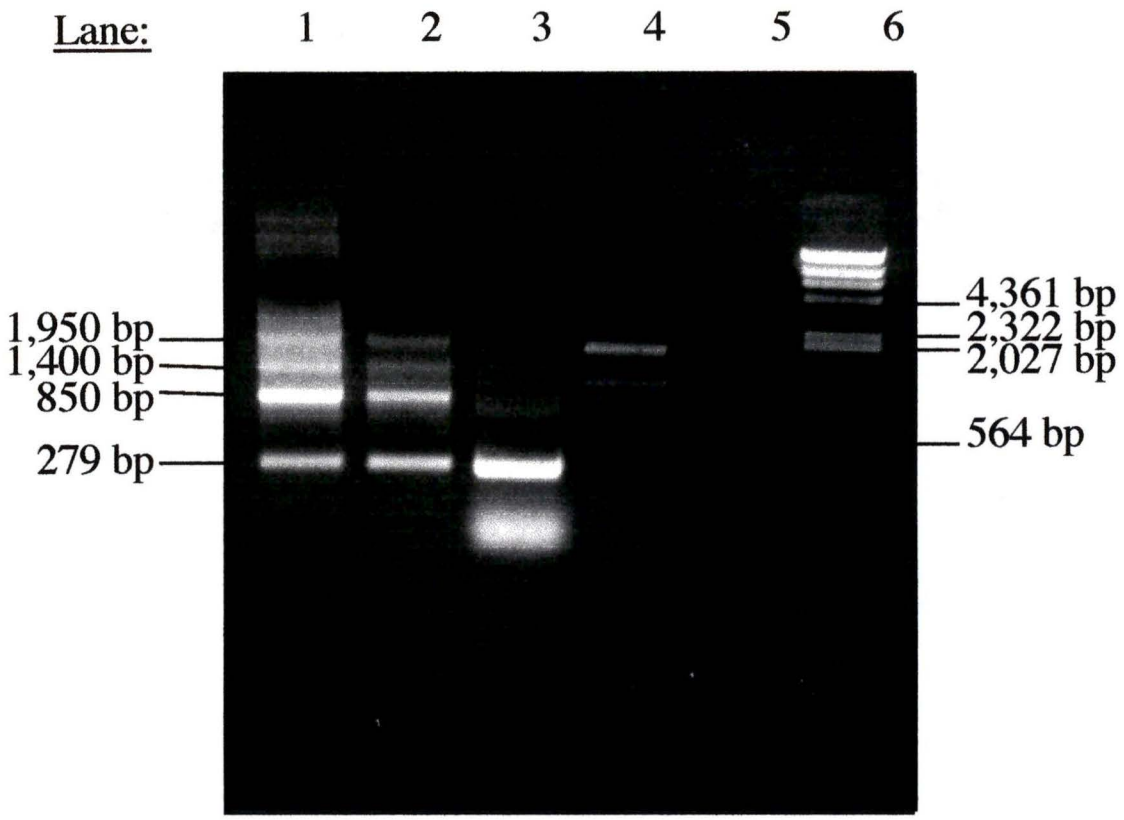


Figure 9: PCR amplification of trypanosome genomic DNA from the P1 bacteriophage clone 3D12. The multiple PCR products observed here were present in repeated experiments. The PCR used forward primer #1 and reverse primer #2 from the KMP-11 gene and the following DNA templates:

Lane 1: pBS KMP-11 2 vector
Lane 2: pBS KMP-11 3c vector
Lane 3: P1 bacteriophage clone 3D12
Lane 4: pBluescript™ plasmid without insert
Lane 5: No template
Lane 6: *Hind* III digested lamda DNA size standards

amplification product was absent in both the of the negative control polymerase chain reactions, which used either the vector pBluescript™ as the template or no template at all (Figure 9, lanes 4 & 5). Further analysis revealed that there were a set of five bands appearing each time the KMP-11 gene was amplified from *T.brucei* genomic DNA using the KMP-11 internal forward and reverse primers. The bands were 300 bp, 850 bp, 900 bp, 1350 bp and 1950 bp long (Figure 10). Other PCR experiments using a combination of either the universal forward or reverse primers of the pBluescript vector plus one of the KMP-11 primers also consistently produced this ladder of bands (data not shown). The PCR data suggested the existence of four KMP-11 genes that were repeated within approximately 2000 bp of one another.

Northern blot analysis of the KMP-11 transcript from a variety of kinetoplastid parasites. RNA was isolated from *Leishmania* promastigotes and procyclic culture forms of several trypanosome strains and the number and relative abundance of the KMP-11 mRNA transcripts were determined. Both *T.b.brucei* and *T.cruzi* showed a set of two relatively abundant KMP-11 transcripts of similar size, while the *L.major* possessed only a single transcript that was larger than that seen for trypanosomes (Figure 11). In the trypanosomes studied, the lower molecular weight KMP-11 transcripts were more abundant than their higher molecular weight counterpart, and the *T.b.brucei* KMP-11 transcripts were more abundant than those from *T.cruzi*. Tubulin transcripts were used as a standard in determining the relative KMP-11 abundance throughout the trypanosome strains tested. The 427.01 strain had the highest KMP-11 transcript levels of the *T.brucei* strains, having 14% and 27% higher levels respectively then the slower growing 348 TI and 348 T2 strains (Figure 11, lanes 1, 2, & 3). The gap in levels between the *T.cruzi* strains was slightly larger, with STIB 745 and STIB 68 MA producing 66% and 70% of the KMP-11 transcript seen in the 1457 strain (Figure 11, lanes 4, 5, & 6). The relatively low levels of KMP-11 transcripts seen for *L.major* may be an artifact related to the use of a heterologous *T.cruzi* KMP-11 probe (Figure 11, lane 7)

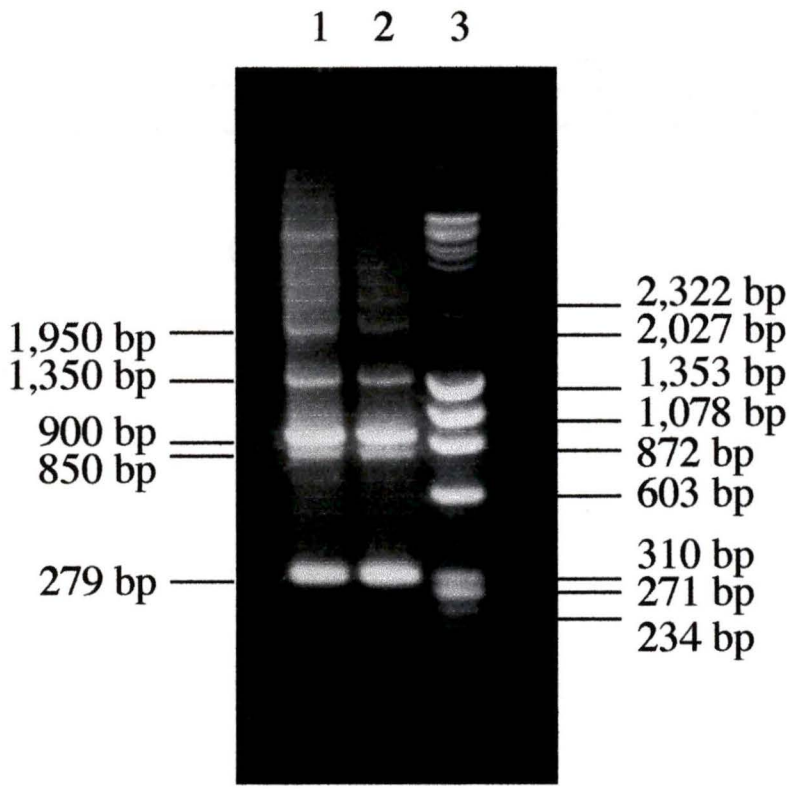
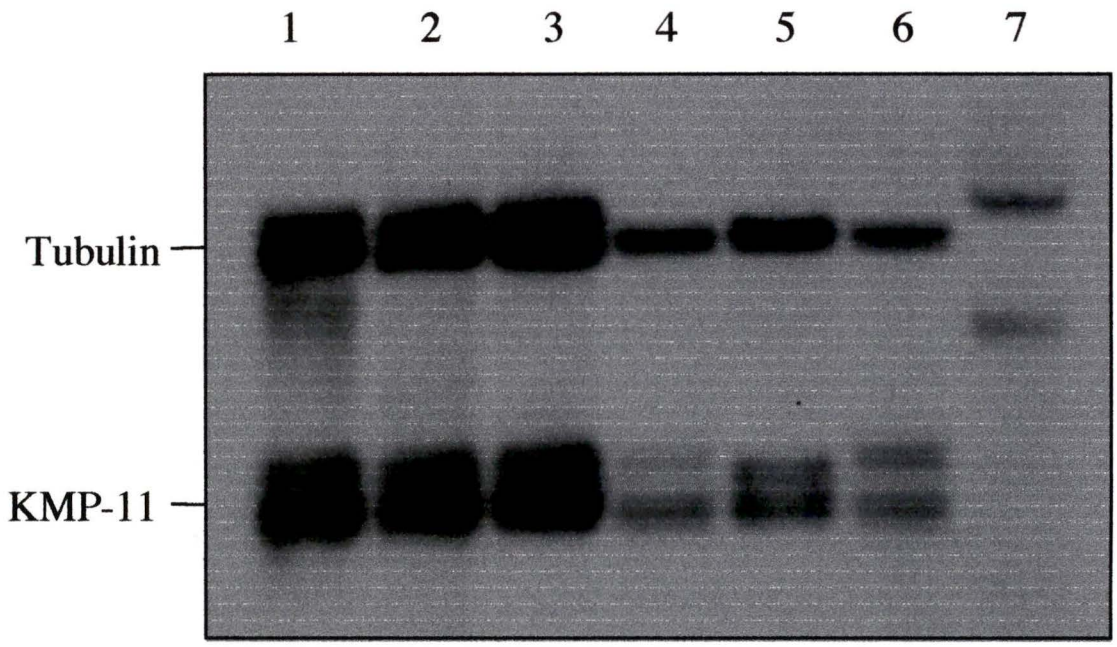


Figure 10: PCR amplification of the KMP-11 gene locus from *T.b.brucei* genomic DNA, using amplification primers #1 and #2. The multiple PCR products observed in lane 1 and 2 in the 1% agarose gel were evident in several similar experiments. Lambda - *Hind* III/*Hea* I DNA size markers are shown in lane 3.



100%	86%	73%	70%	66%	100%

Figure 11: Northern blot analysis of mRNA from a variety of log-phase kinetoplastid parasites. The blot was hybridized simultaneously with a KMP-11 gene probe and the alpha-tubulin open reading frame from *T.cruzi*, both of which were ^{32}P-labeled. Below the figure the relative abundance of KMP-11 transcripts are given.

Lane 1: mRNA from *T.b.brucei* 427.01
Lane 2: mRNA from *T.b.brucei* 348 TI
Lane 3: mRNA from *T.b.brucei* 348 T2
Lane 4: mRNA from *T.cruzi* STIB 68 MA
Lane 5: mRNA from *T.cruzi* STIB 745
Lane 6: mRNA from *T.cruzi* 1457
Lane 7: mRNA from *L.major* YS119

Due to the apparent complexity of the KMP-11 locus elucidated in the previously discussed PCR experiments, sequencing of the region was performed using the 3500 bp trypanosome genomic DNA insert that had been subcloned into the pBS KMP-11 3c construct.

Generation and sequencing of the KMP-11 "shotgun" sequencing library. A large amount of the pBS KMP-11 3c insert was isolated after its PCR amplification using the forward and reverse universal pBluescript primers and the fragment was subsequently nebulized. The nebulization was halted after 30 seconds and an aliquot of the DNA was removed and run on an agarose gel to determine the extent of disruption (Figure 12, part A). At this time point the 3500 bp insert was largely intact and only a slight DNA smear at lower molecular weights was visible. This procedure was repeated for 60 seconds producing a much wider smear that showed the highest concentration of bands between 1500 bp and 2000 bp (Figure 12, part B). After a third and final nebulization, the total elapsed time was 120 seconds and the 3500 bp insert had been transformed into a myriad of bands ranging from approximately 3000 bp to less than 100 bp (Figure 12, part C). The bands between 1000 bp and 750 bp were relatively abundant and were an ideal insert size for sequencing. Thus bands of this range were subsequently isolated from an agarose gel (Figure 13) and cloned into the M13 bacteriophage to generate the sequencing library. Over 120 plaques from the M13 bacteriophage library were directly PCR tested to obtain the final 72 clones that were sequenced. The direct PCR testing used the M13 universal forward and reverse primers that flanked the cloning site, and allowed the quick and easy identification of usable clones (Figure 14). After 72 individual reads of approximately 550 bp on the automated DNA sequencer there was enough data to assemble a comprehensive sequence of the KMP-11 locus, having an 11.5 fold sequence redundancy.

Sequence of the KMP-11 locus. The sequence assembled using the Seqman™ sequence alignment software (DNASTAR, Madison, WI, USA), was 3375 bp in length

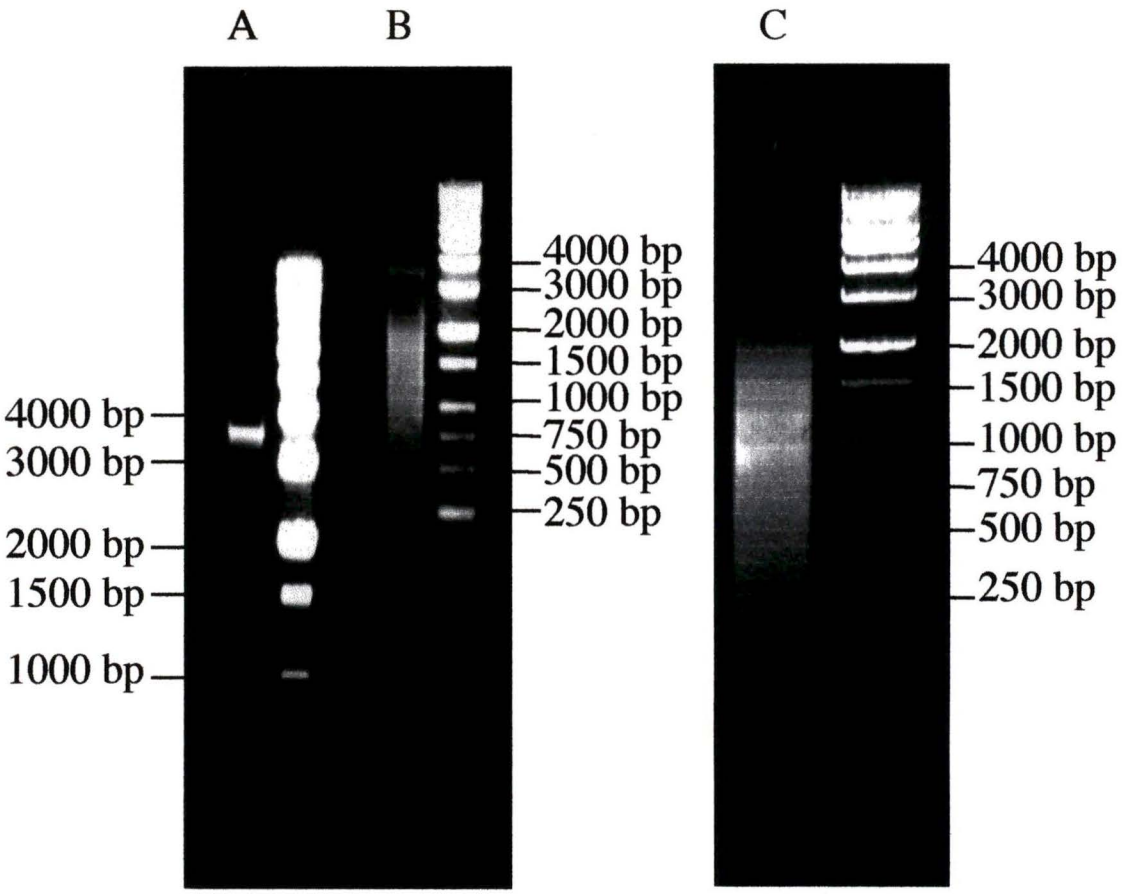


Figure 12: Nebulized KMP-11 insert DNA after (A) 30 seconds, (B) 60 seconds and (C) 120 seconds of high pressure disruption. In each case a Kb DNA ladder standard was run along side the nebulized DNA.

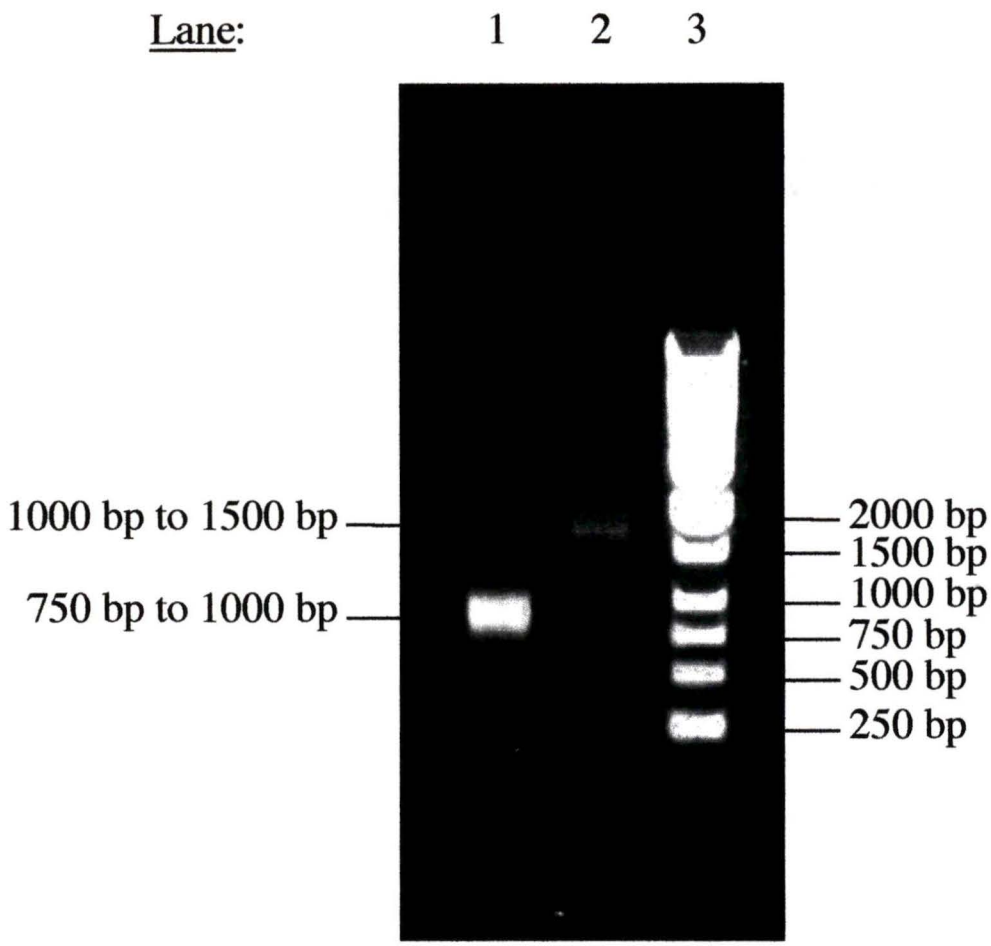
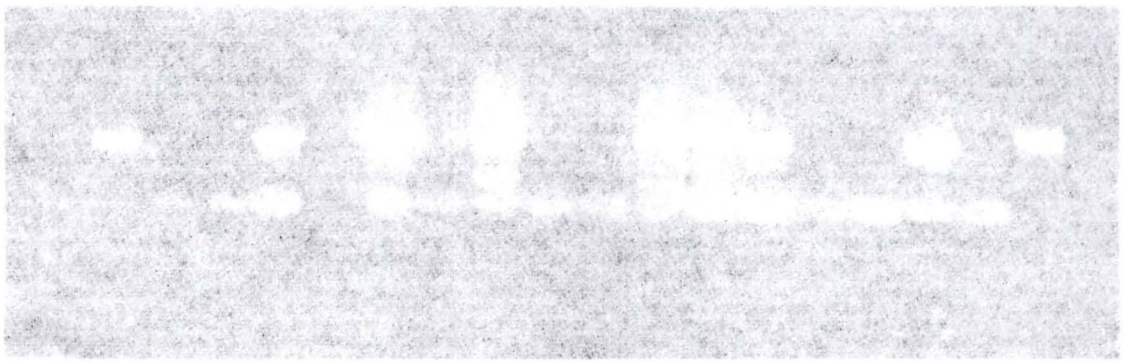


Figure 13: Nebulized KMP-11 genomic DNA fragments used in the production of the M13 bacteriophage sequencing library. The DNA fragments from the 750 bp to 1000 bp band (Lane 1) were used for sequencing as the fragments in the 1000 bp to 1500 bp band (Lane 2) were a less manageable sequencing size. The Kb DNA ladder standards were run alongside the nebulized genomic DNA (Lane 3).

Figure 14: PCR amplification of M13 bacteriophage plaques produced during synthesis of the KMP-11 sequencing library. Inserts were amplified using the M13 universal forward and reverse primers flanking the bacteriophage cloning site. The positive clones (72, 74, 77, 78, 79, 82, 84, 86, 89) containing inserts were among the clones used to generate a comprehensive KMP-11 gene locus sequence. A negative control PCR ("-") was loaded in the lane beside clone 89.



and contained four KMP-11 genes (Figure 15). The sequence contained 704 bp of upstream-untranslated region and 632 bp of downstream untranslated flanking sequence. The four KMP-11 genes start 705 bp, 1298 bp, 1871, and 2464 bp into the sequence, leaving only small uniform intergenic regions of 320 bp, 300 bp and 320 bp between the four genes. The sequenced KMP-11 locus was A-T rich containing long polypyrimidine stretches in front of each KMP-11 gene, in all containing 55.5% adenosine/thymine residues and 44.5% guanine/cytosine residues.

Dot matrix analysis for elucidation of repeated regions within the KMP-11 locus. The dot matrix analysis plots regions of similarity between two DNA sequences where a co-linear region appears as a diagonal line of dots. In a dot matrix comparison of the KMP-11 sequence against itself, diagonal lines other than the central diagonal reveal four repeats (K1-K4) of approximately 380 bp that are 98.7% to 100% identical (Figure 16). Each of the repeats contains a KMP-11 gene and 102 bp of upstream untranslated sequence, with the exception of the K3 repeat which contains only 85 bp of upstream repeat due to an apparent 17bp deletion. Two larger repeats, R1 (973 bp) and R2 (957 bp) are also evident from the dot matrix analysis. R1 encompasses K1, K2 and the intergenic sequence between them, while R2 includes K3, K4 and their intergenic region. Within this entire stretch the only non-duplicated portion is the intergenic region of 209 nucleotides between K2 and K3.

Detailed analysis of the four KMP-11 repeats. Due to the high level of sequence repetition and the similarity of the repeated gene sequences (98.7% - 100%) careful and comprehensive analysis of the sequence data was required. Within the repeated KMP-11 genes there are only two positions at which the sequences diverge (Figure 17). At a position 9 nucleotides from the start of the gene, the fourth KMP-11 repeat (K4) has undergone a cytosine to thymine transition, while at a position 4 nucleotides from the end

```
   1 tccgttgcaa ggtttggctc tgcggtggtg ctgcggtcgg ttgccgacct cttcttcgat
  61 gttgctcttg cgattcggca tctgacgatg ggggaaatgt tgtcattgcg tcgtgtatgt
 121 gatgatgaag cgctcattac tgttcccaga gcagttggaa tccggcttct tgaactagct
 181 tcagagggtc acggccttca agtaactgtc gcggttggtg aagaagaaca aaacggttgg
 241 aaggttaccg aacaatctca gtgcgttagc tattatttat ccgtgcacaa accccgtgcg
 301 caggtgatgg tggatggtga gaacctacaa ttggtgaaga agcatgaggt attgcgcagc
 361 atggcaaaca tgatggaaca caccaacgca aaagtggcag gattcgtcga cccgtttggt
 421 aatgtcctta ttcctagaca atgaccagga cggggagcgc gacaaggtta agtgatggtt
 481 tcttccacca tgtcaccact ggtgcgggag cgacgtgcca attcatttga ttccccccct
 541 tacatttgtg cagtcatttg tttatcctag aaccaacttc tacctgtaac acttttccc
 601 ctttatcgtt tttcttttct ttcctttcca ttacgccact tttcatttcg tttagatcgt
 661 gctttcaaga gtaaactaaa ctttaagata ctttttcaag aaacatggcc accacatacg
 721 aagaatttgc tgcgaagctc gaccgcctcg atgccgaatt cgccaagaag atggaggagc
 781 agaacaagcg attcttcgct gacaagcctg atgaggctac gctgtcccct gagatgaaag
 841 agcactatga aaagttcgaa aaaatgatcc aggagcacac ggacaagttc aacaagaaga
 901 tgcgcgagca ctcagagcac ttcaaggcca agtttgcgga actcctcgag cagcagaaga
 961 atgcccagtt ccccggaaag tgatttcgaa aacactctcc ttttgtttct tttctctctt
1021 attagaggtc cttccgggca ctatgttact aaattccacc tctttggtgt ctttgcaagg
1081 gggctacctc tgtcaaacta ttgatatctc taaaataacg tcccggatca tgtgttgccg
1141 tggcgcagtt tttacttacc tcttccacca cttgtactgc cacttgcttg acctatatcg
1201 ttttttcttt tctttccttt ccgtttcgct acttttcatt tcgtttagat tgtgctttca
1261 agagtaaact aaactttaag atacttttc aagaaacatg gccaccacat acgaagaatt
1321 tgctgcgaag ctcgaccgcc tcgatgccga attcgccaag aagatggagg agcagaacaa
1381 gcgattcttc gctgacaagc ctgatgaggc tacgctgtcc cctgagatga aagagcacta
1441 tgaaaagttc gaaaaaatga tccaggagca cacggacaag ttcaacaaga agatgcgcga
1501 gcactcagag cacttcaagg ccaagtttgc ggaactcctc gagcagcaga agaatgccca
1561 gttccccgga aaatgattac aaatttacta cacctctcat tttctttcct atgttagagt
1621 ccctagggat aatactataa aaatctttta aaagatgtaa ataccacttt tctagcgcaa
1681 tgcatgagta tatgaattaa tattggttgc cactgttaac ttacttgtcc cgttcacttc
1741 ggtgtgcctc cctatccctt atccctgaat aagttgctcc tccttttcc tttccgttac
1801 gccacttttc atttcgttta gatcgtgctt tcaagagtaa actaaacttt aagatacttt
1861 ttcaagaaac atggccacca catacgaaga atttgctgcg aagctcgacc gcctcgatgc
1921 cgaattcgcc aagaagatgg aggagcagaa caagcgattc ttcgctgaca agcctgatga
1981 ggctacgctg tcccctgaga tgaaagagca ctatgaaaag ttcgaaaaaa tgatccagga
2041 gcacacggac aagttcaaca agaagatgcg cgagcactca gagcacttca aggccaagtt
2101 tgcggaactc ctcgagcagc agaagaatgc ccagttcccc ggaaagtgat ttcgaaaaca
2161 ctctcctttt gtttcttttc tctcttatta gaggtccttc cgggcactat gttactaaat
2221 tccacctctt tggtgtcttt gcaaggggc tacctctgtc aaactattga tatctctaaa
2281 ataacgtccc ggatcatgtg ttgccgtggc gcagttttta cttacctctt ccaccacttg
2341 tactgccact tgcttgacct atatcgtttt ttcttttctt tcctttccgt ttcgctactt
2401 ttcatttcgt ttagattgtg ctttcaagag taaactaaac tttaagatac tttttcaaga
2461 aacatggcca ctacatacga agaatttgct gcgaagctcg accgcctcga tgccgaattc
2521 gccaagaaga tggaggagca gaacaagcga ttcttcgctg acaagcctga tgaggctacg
2581 ctgtcccctg agatgaaaga gcactatgaa aagttcgaaa aaatgatcca ggagcacacg
2641 gacaagttca acaagaagat gcgcgagcac tcagagcact tcaaggccaa gtttgcggaa
2701 ctcctcgagc agcagaagaa tgcccagttc cccggaaaat gatgatagaa catatgtatt
2761 taaggagact tctaaccccg tgaggatgtc acatattttt ttttcgttga ggtcttgatg
2821 aatttctatt gaatataatt ttttttgatc ttttttgagt tcctttacta ctaccccaag
2881 caacggatgt tttttttgtg tagtgggagg gacgaatatg tttgaactct ctgcaaatga
2941 atagaatcta tttttccttg tactcagcac ctacttgaag gaatttatga ttgtggttaa
3001 gctgacactc atctgtgcat atttttgga cgcgtatgta tcttacttga ttatcctttc
3061 cctccattct ctgtaatcgt attttcctct ggttctgcac tgtaaatagc ggtagtaacc
3121 tcgaagtgta gggaaacttt gaaggggact actgatttga tttttctact ttctgcggcc
3181 gttgcgtctt agtgtgtgtt tggtcatata gcgtgtattc ccttgtggta gtgtttttt
3241 ttttttggc agagcgggga gaatgctgtg ccgtaccttt ttacggcaat tcaggatgag
3301 tggtggggat atgtttgtgg agtataaagt gctttcccgt gaccatagac ggtcgattcg
3361 tgttgaggac gcgat
```

Figure 15: Sequence of the 3500bp *Bam*HI fragment from the *T.b.brucei* strain TREU/4927 P1 bacteriophage clone which contained the KMP-11 gene locus. The four open reading frames of the KMP-11 gene copies are underlined.

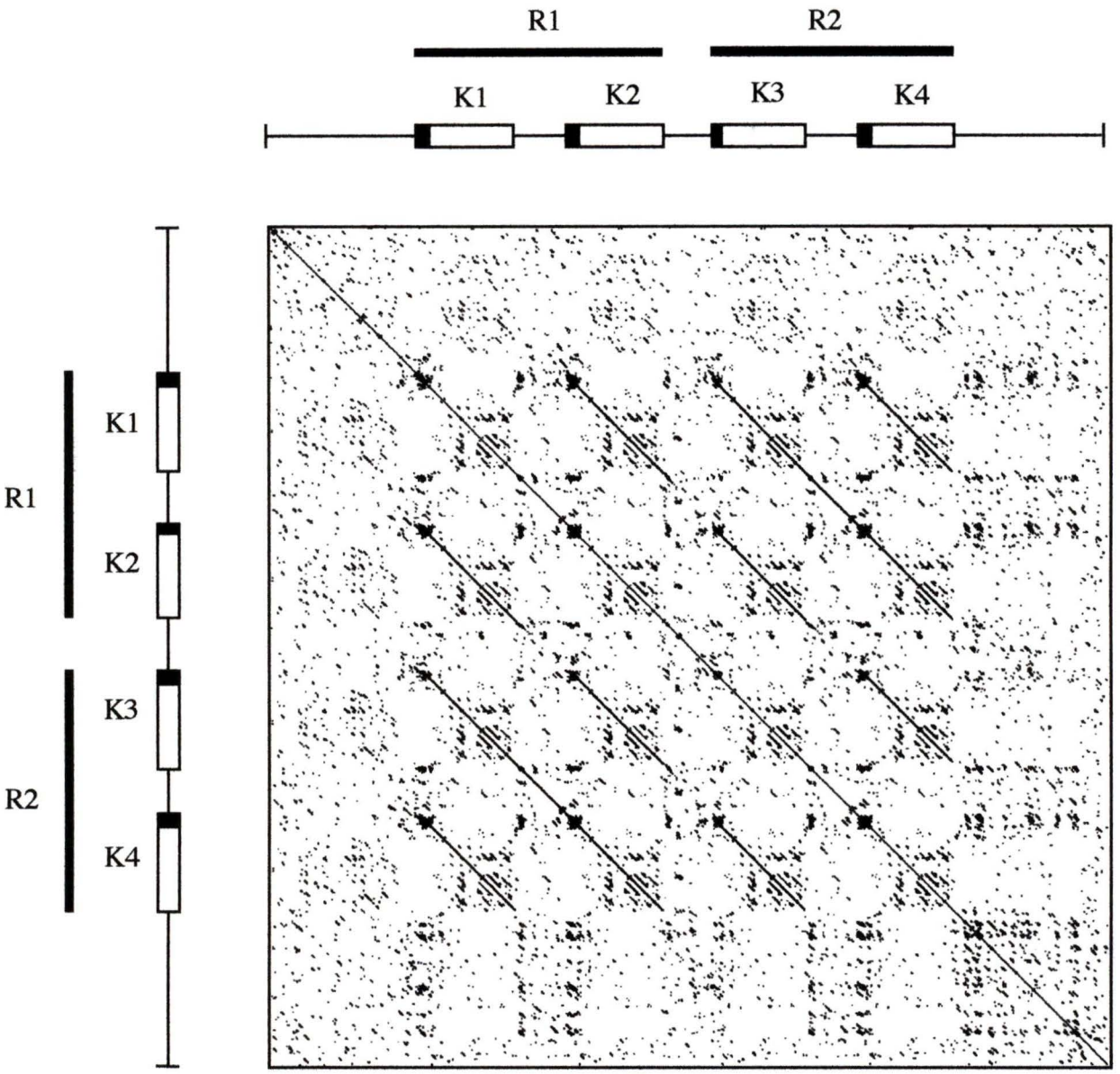


Figure 16: Dot matrix analysis of the KMP-11 locus DNA sequence. Excluding the main diagonal, diagonal lines indicate regions of internal sequence similarity. Along the axes, the open rectangles designate KMP-11 genes while the closed boxes designate conserved 5'UTR sequences. Together these symbols represent the smallest KMP-11 gene duplication unit. The long closed rectangles, R1 and R2, indicate larger regions of repeated DNA. The Geneworks™ dot matrix analysis program (IntelliGenetics Inc., Mountain View CA) was used to generate the plot where a dot corresponds to a window size of 20 bp with a minimum number of 14 matches within that window and an offset of 10 bp.

```
                                                  *   *   *
K1          TATCGTTTTT TCTTTTCTTT CCTTTCCATT ACGCCACTTT TCATTTCGTT          50
K2          TATCGTTTTT TCTTTTCTTT CCTTTCCGTT TCGCTACTTT TCATTTCGTT          50
K3          ---------- -------TTT CCTTTCCGTT ACGCCACTTT TCATTTCGTT          33
K4          TATCGTTTTT TTCTTTCTTT CCTTTCCGTT TCGCTACTTT TCATTTCGTT          50

               *
K1          TAGATCGTGC TTTCAAGAGT AAACTAAACT TTAAGATACT TTTTCAAGAA         100
K2          TAGATTGTGC TTTCAAGAGT AAACTAAACT TTAAGATACT TTTTCAAGAA         100
K3          TAGATCGTGC TTTCAAGAGT AAACTAAACT TTAAGATACT TTTTCAAGAA          83
K4          TAGATTGTGC TTTCAAGAGT AAACTAAACT TTAAGATACT TTTTCAAGAA         100

                     *
K1          ACATGGCCAC CACATACGAA GAATTTGCTG CGAAGCTCGA CCGCCTCGAT         150
K2          ACATGGCCAC CACATACGAA GAATTTGCTG CGAAGCTCGA CCGCCTCGAT         150
K3          ACATGGCCAC CACATACGAA GAATTTGCTG CGAAGCTCGA CCGCCTCGAT         133
K4          ACATGGCCAC TACATACGAA GAATTTGCTG CGAAGCTCGA CCGCCTCGAT         150
               M  A   T  T  Y  E   E  F  A   A  K  L   D  R  L  D

K1          GCCGAATTCG CCAAGAAGAT GGAGGAGCAG AACAAGCGAT TCTTCGCTGA         200
K2          GCCGAATTCG CCAAGAAGAT GGAGGAGCAG AACAAGCGAT TCTTCGCTGA         200
K3          GCCGAATTCG CCAAGAAGAT GGAGGAGCAG AACAAGCGAT TCTTCGCTGA         183
K4          GCCGAATTCG CCAAGAAGAT GGAGGAGCAG AACAAGCGAT TCTTCGCTGA         200
              A  E  F   A  K  K   M  E  E  Q   N  K  R   F  F  A

K1          CAAGCCTGAT GAGGCTACGC TGTCCCCTGA GATGAAAGAG CACTATGAAA         250
K2          CAAGCCTGAT GAGGCTACGC TGTCCCCTGA GATGAAAGAG CACTATGAAA         250
K3          CAAGCCTGAT GAGGCTACGC TGTCCCCTGA GATGAAAGAG CACTATGAAA         233
K4          CAAGCCTGAT GAGGCTACGC TGTCCCCTGA GATGAAAGAG CACTATGAAA         250
            D  K  P  D   E  A  T   L  S  P   E  M  K  E   H  Y  E

K1          AGTTCGAAAA AATGATCCAG GAGCACACGG ACAAGTTCAA CAAGAAGATG         300
K2          AGTTCGAAAA AATGATCCAG GAGCACACGG ACAAGTTCAA CAAGAAGATG         300
K3          AGTTCGAAAA AATGATCCAG GAGCACACGG ACAAGTTCAA CAAGAAGATG         283
K4          AGTTCGAAAA AATGATCCAG GAGCACACGG ACAAGTTCAA CAAGAAGATG         300
             K  F  E   K  M  I  Q   E  H  T   D  K  F   N  K  K  M

K1          CGCGAGCACT CAGAGCACTT CAAGGCCAAG TTTGCGGAAC TCCTCGAGCA         350
K2          CGCGAGCACT CAGAGCACTT CAAGGCCAAG TTTGCGGAAC TCCTCGAGCA         350
K3          CGCGAGCACT CAGAGCACTT CAAGGCCAAG TTTGCGGAAC TCCTCGAGCA         333
K4          CGCGAGCACT CAGAGCACTT CAAGGCCAAG TTTGCGGAAC TCCTCGAGCA         350
              R  E  H   S  E  H   F  K  A  K   F  A  E   L  L  E

                                                *
K1          GCAGAAGAAT GCCCAGTTCC CCGGAAAGTG A                            381
K2          GCAGAAGAAT GCCCAGTTCC CCGGAAAATG A                            381
K3          GCAGAAGAAT GCCCAGTTCC CCGGAAAGTG A                            364
K4          GCAGAAGAAT GCCCAGTTCC CCGGAAAATG A                            381
            Q  Q  K  N   A  Q  F   P  G  K
```

Figure17: Alignment of the four KMP-11 repeats (K1-K4) found within the sequence of the KMP-11 locus using the Geneworks™ DNA alignment program. Each repeat contains a complete KMP-11 gene and conserved upstream untranslated sequence. The predicted amino acid sequences of the four genes are shown under the nucleotide sequences. Nucleotide mismatches are indicated by '*'. Deleted sequence is identified by '-'. Pyrimidine-rich elements are underlined and putative splice acceptor sites are in bold type. The start and stop codons for the KMP-11 genes are shaded.

of the gene sequence, both the second (K2) and the fourth (K4) KMP-11 genes have undergone a guanine to adenine transition. In both cases these changes have occurred in the third (wobble) position of the codon, and do not alter the predicted amino acid sequence of the protein (Figure 17). The upstream 5' untranslated region (UTR) is less well conserved with four points of sequence divergence at residues 47 bp, 68 bp, 72 bp and 75 bp upstream of the KMP-11 start codon, as well as a deleted stretch of 17 bp from the most 5' end of the K3 repeat. Long polypyrimidine rich elements are evident in the 5' UTR which may play a role in the transcriptional and post-transcriptional events. Putative splice acceptor sites are indicated in bold type. While the sequence conservation extends to the 5'UTR it ends with the stop codon of the KMP-11 genes leaving three of the four 3' UTR with little or no sequence similarity. The sequence similarity between the 3' UTR of the K1 and K3 repeats was likely the result of a duplication event that generated the R2 repeat from the R1 region.

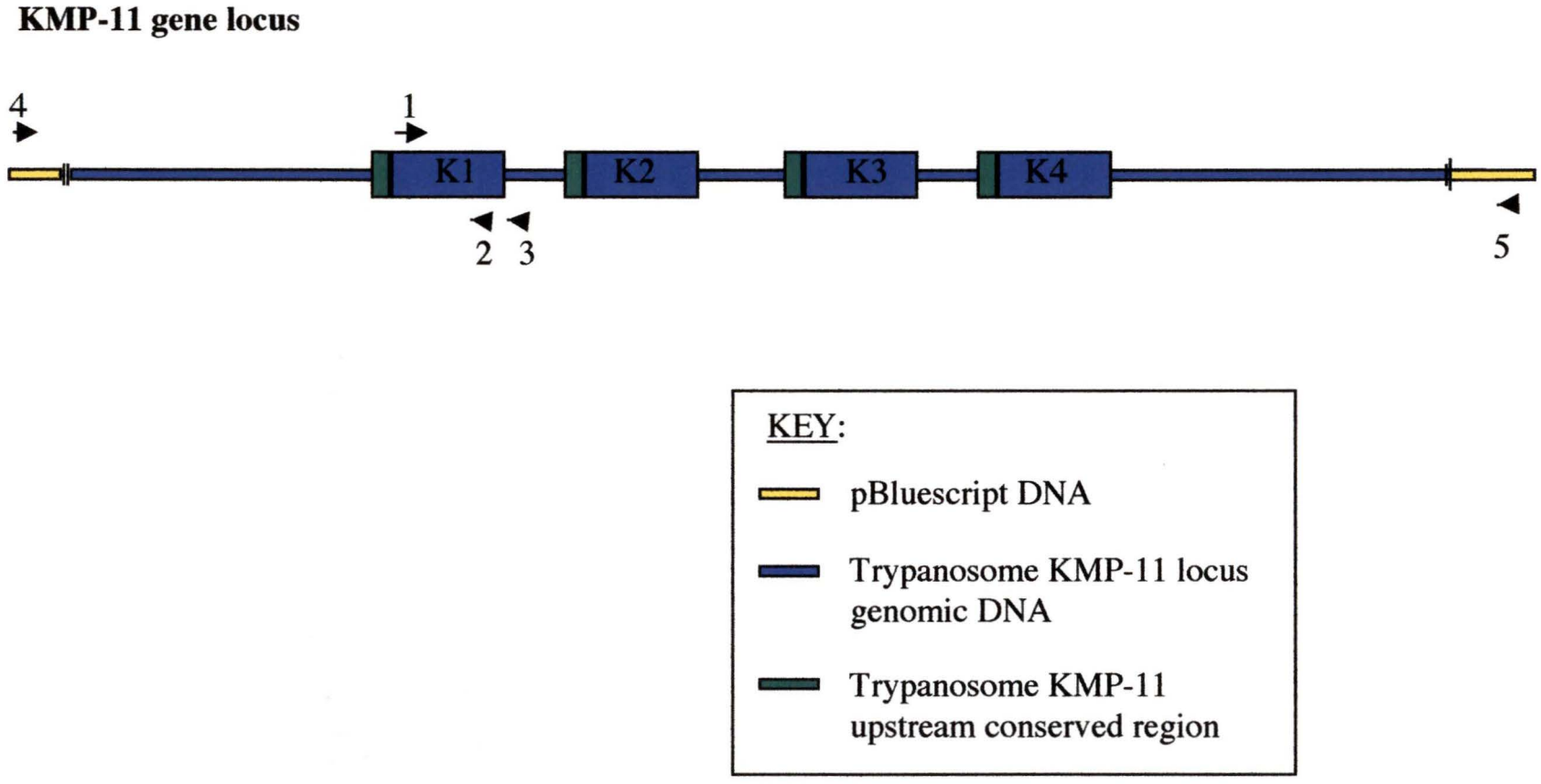


Figure 18: Schematic of the KMP-11 gene locus and primers generated for KMP-11 locus sequencing.

Primer Number	Primer Name	Primer use	Primer Sequence
Primer #1	KMP-11 gene amplification forward	Southerns, Diagnostic PCR	5'-ATG GCC ACC ACA TAC GAA G-3'
Primer #2	KMP-11 gene amplification reverse (A)	Southerns, Diagnostic PCR	5'-TCA TTT TCC GGG GAA CTG-3'
Primer #3	KMP-11 gene amplification reverse (B -external)	Southerns, Diagnostic PCR	5'-AAT GGA AAG AAA ATG AGA GGT G-3'
Primer #4	pBS forward universal primer (-20)	PCR, Shot gun Sequencing	5'-GTA AAA CGA CGG CCA GT-3'
Primer #5	pBS reverse universal primer (M13)	PCR, Shot gun Sequencing	5'-CAG GAA ACA GCT ATG AC-3'

Table 1: Polymerase chain reaction primers used during the KMP-11 gene locus sequencing experiments.

DISCUSSION

The trypanosome genome contains many examples of protein-encoding genes that are arranged in tandem repeats separated by unrelated intergenic sequences. These tandem arrays are often polycistronically transcribed and subsequently undergo posttranscriptional modification to generate mature monocistronic mRNAs. Accurate 5' cap addition and 3' polyadenylation of the mRNA is dependent on at least four elements including pyrimidine-rich stretches within the intergenic region [Schurch *et al.*, 1994]. The sequence data reported here for the KMP-11 locus shows several of these features. The region just upstream of each of the KMP-11 genes is highly conserved and shows pyrimidine rich tracts, as well as several potential splice acceptor sites (Figure 17). All of the intergenic regions within the KMP-11 array have putative polyadenylation sites 50 to 100 nucleotides upstream of the polypyrimidine elements which suggests that all four genes may contribute to KMP-11 protein production *in vivo*. It has been suggested that the control of gene expression operates primarily at a posttranscriptional level due to the trypanosome polycistronic transcription system [Vanhamme and Pays, 1995]. Imperfect inverted repeats positioned near the beginning of each KMP-11 gene may contribute to this regulation through the generation of hairpin loops. These hairpin loops may alter the ability of one or more of these linked genes to be translated or may alter RNA stability. Trypanosome bloodstream forms and procyclic culture forms differentially express this 11-kDa protein [Stebeck *et al.*, 1995], but whether this regulation is due to an inhibitory hairpin loop in the mRNA or occurs at a later stage of protein production is still unresolved.

The complete KMP-11 locus DNA sequence gives us insight into the evolutionary duplication events that likely occurred to produce this multiple gene copy locus. Presumably, the initial duplication of the KMP-11 gene involved the coding region and the first 102 nucleotides of the 5' untranslated region (UTR). This 5' UTR that contains pyrimidine tracts that are essential for accurate posttranscriptional modifications must have had an important role in the expression of the primary KMP-11 molecule to necessitate its duplication. The new gene was positioned 314 nucleotides away from its sister, just

downstream from an element that could act as a polyadenylation site. This process generated what we now see as the large repeat R1 (Figure 16). In the intervening period, the aforementioned guanine to adenine transition altered the sequence of the second KMP-11 gene. Both KMP-11 genes, the conserved 5' UTR and the functional intergenic region were then duplicated and translocated to a position 209 bp downstream of the original repeat. During or after this process there was a deletion of 17 bp from the 5' end of the new repeat, R2, and a transition within the fourth KMP-11 gene, producing the present day KMP-11 locus observed in this *T.brucei* clone.

Characterization of the KMP-11 locus was an essential preliminary step in our study of the biological function of KMP-11. Four highly conserved copies of the KMP-11 gene were identified within a region of 2200 bp. The tightly packed nature of the KMP-11 locus should allow deletion studies to proceed through a single step knockout procedure. The presence of two allelic KMP-11 loci in the diploid trypanosome made it necessary to perform two rounds of loci deletions to generate null/null KMP-11 mutants. One KMP-11 loci was knocked out using the neomycin drug resistance gene marker and the other loci was knocked out using the phleomycin drug resistance marker in hopes of facilitating structure-function studies of the KMP-11 protein.

CHAPTER 2: Attempted Genetic Deletion of the Kinetoplastid Membrane protein-11 Protein of *Trypanosoma Brucei*

INTRODUCTION

African trypanosomes exhibit many novel biochemical control mechanisms such as antigenic variation, polycistronic transcription, RNA editing and digenetic regulation of vast arrays of genes between life cycle stages. A system of stable transformation allowing targeting of exogenous homologous DNA to specific genetic loci has recently been established in a variety of kinetoplastids and may facilitate functional studies of genes and gene systems. To date, these stable transformations have been used to study a variety of gene targets in *Leishmania* [Kelly *et al.*, 1992; Webb *et al.*, 1994], *T. brucei* [ten Asbroek *et al.*, 1990, 1993; Lee *et al.*, 1990, 1991; Eid and Webb, 1991a, 1991b; Jefferies *et al.*, 1992; and Blundell *et al.*, 1996] and *T.cruzi* [Kelly *et al.*, 1992]. The stable transformation involves homologous recombination of highly related DNA sequences that direct the integrative events. Similar systems have been characterized in bacteria, yeast and mammals but have differing efficiencies and requirements than the trypanosome system. In mouse embryo-derived stem cells very low levels of targeted transformants were generated by homologous recombination with 3 Kb of homologous DNA while 20 fold higher levels were seen when four fold as much homologous DNA was used [Deng and Capecchi, 1992]. Ten to one hundred times less DNA is require for homology-mediated targeting in *E.coli* and yeast as less than 100 bp of homologous DNA is essential [Manivasakam *et al.*, 1995; Shen *et al.*, 1986]. Trypanosomes and yeast both integrate exogenous DNA into the genome essentially entirely by homologous recombination [Blundell *et al.*, 1996; Orr-Weaver *et al.*, 1981], while mammalian cells show integration by this means in less than 1% of the transformants assayed [Jasin *et al.*, 1990].

The requirement for precise homology of target and donor DNA has been shown by the suppression of recombination between divergent sequences found within the trypanosome

VSG expression sites. Inserted DNA is targeted almost exclusively to the expression site from which it has been derived [Brundell *et al.*, 1996]. When sites that were 92% identical were examined, it was shown that all the insert DNA had targeted into the homologous site and not into the heterologous site that only differed by 8%. This data suggests that the trypanosome may lack a mechanism allowing random insertion of exogenous DNA. An interesting paradox thus arises with respect to the method of recombination that occurs during antigenic variation, one of the most important parasite immune evasion mechanisms. With antigenic variation of VSG, recombination occurs despite the fact that the incoming VSG gene has limited sequence identity to the gene already present in the expression site. Short stretches of upstream and downstream homology exist as 5' imperfect 70 bp repeats and 3' telomeric repeats, respectively [Vanhamme and Pays, 1995]. In any given recombination event it is suggested that the largest blocks of homology would be 70 bp at the 5' end and 14 bp at the 3' end of the VSG genes involved. This use of short and somewhat dissimilar repeats for targeting of VSG homologous recombination suggests that trypanosomes may either accept low efficiencies of transposition of the VSG genes or use alternative methods for VSG gene recombination. In the alternative methods there is likely suppression of the mismatch repair systems which is thought to prevent imprecise homologous integrations [Brundell *et al.*, 1996]. Indeed recent results suggest that differentially regulated recombination occurs in the VSG site depending on the time within a given stage [Vanhamme and Pays, 1995].

The requirements of homologous integration events in trypanosomes have been partially defined in recombination experiments involving the targeted disruption of the calmodulin locus [Eid and Webb, 1991]. The transformants showed a variety of integration patterns that allowed differences between the first and second crossover events to be distinguished. The initial recombinational event of the *T.brucei* double crossover appears to be similar to that in yeast and mammals in which the efficiency is dependent on the length of the homologous DNA. The second crossover event seems to be less dependent on the total length of the homologous sequence, because few of the calmodulin directed transformants used the entire stretch of homologous sequence. Instead, most of the

crossovers occurred in a small segment of DNA closest to the neomycin drug resistance gene carried by the disruption cassette. The second crossover event thus very often occurs in the most proximal homologous DNA to the DNA integrated in the first crossover event [Eid and Webb, 1991].

The targeted deletion cassettes used for knockout mutagenesis in trypanosomes possess one of several drug resistance genes that allow positive selection of transformants. Two of the most frequently used of these trypanosome compatible markers is the neomycin phosphotransferase gene [ten Asbroek *et al.*, 1990], and the phleomycin (sh *ble*) resistance gene [Jefferies *et al.*, 1993]. The existence of multiple trypanosome drug resistance markers is very useful for the disruption of both alleles of a gene in a diploid organism. Neomycin is an aminoglycoside related to kanamycin and gentamycin, which binds ribosomes and inhibits subsequent protein synthesis. Neomycin phosphotransferase phosphorylates the drug inactivating it by interference with its active transport into the cell [Davis and Smith, 1979]. The phleomycin antibiotic which is related to bleomycins, is a glycopeptide that is active at low concentration and binds and cleaves DNA [Drocourt *et al.*, 1990]. The phleomycin resistance gene (sh *ble*) encodes a small acidic protein which binds strongly to the phleomycin antibiotic, stopping its activation by ferrous ions and oxygen. The Sh *ble* protein sequesters phleomycin in a one to one ratio, physically inhibiting the breakdown of the DNA [Galignol *et al.*, 1988].

In the current study, targeted gene deletions of both allelic copies of the KMP-11 locus were undertaken in an attempt to determine the functional relevance of KMP-11 in African trypanosomes. For this work a set of deletion cassettes containing the phleomycin or neomycin drug resistance genes, flanked by homologous DNA from the KMP-11 locus were constructed. The DNA upstream of the first KMP-11 gene and downstream of the fourth KMP-11 gene found in the gene locus was integrated into the cassette. The inclusion of homologous DNA from regions flanking the KMP-11 locus should direct the exogenous deletion cassette into the genomic KMP-11 locus.

MATERIALS AND METHODS

Trypanosomes. T.b.brucei 427.01 and *T.b.brucei* 348 PCF were maintained in procyclic culture medium at 27° C. Cultures were generally initiated a week prior to their use in electroporation experiments and split the day before the experiment. After transformation experiments and cloning of transformed trypanosomes by limiting dilution, the newly isolated clones (at approximately 5 x10^6 trypanosomes/ml), were transferred to freezing medium (procyclic culture medium containing 20% FBS, 20% glycerol and the appropriate antibiotics) and stored in liquid nitrogen. When required, the frozen stocks were removed from the liquid nitrogen, warmed briefly at 37° C until thawed and transferred to a 30 ml tissue culture flask containing medium (plus 20% FBS), that had been pre-gassed with a 5% CO_2 in air mixture. Cultures were incubated at 27° C.

Generation of the KMP-11 locus deletion cassettes. Tetracycline resistant XL-1 Blue *E.coli* cells (Stratagene, La Jolla, CA, USA) were used in high voltage electrophoretic transformation experiments according to the method of Dower et al [1988]. Standard gene cloning techniques described by Sambrook [Sambrook *et al.*, 1989] were used. Bacterial cultures were grown in Luria-Beitani (LB) medium containing 1% bacto-tryptone, 0.5% bacto-yeast extract and 0.17 M NaCl, pH 7.0 [Sambrook *et al.*, 1989]. DNA inserts were recovered from low-melting temperature agarose using Qiaex II gel extraction Kit (Qiagen, Chatsworth, CA, USA) and cloned into the pBluescript™ (pBS) neo or pBS phleo plasmid constructs that were provided by Dr. Isabel Roditi (Institut fur Allgemeine Mikrobiologie, Universitat Bern, Bern, SW). pBS neo was derived from the pSV2 neo plasmid and contained the neomycin drug resistance gene inserted into a pBS plasmid backbone using the *Hind* III and *Bam*HI sites of the multiple cloning site. The pBS phleo construct was derived from the pHD63 vector and contained a phleomycin drug resistance gene similarly inserted into *Hind* III and *Bam*HI endonuclease restriction sites. The 5' untranslated region

upstream of the first KMP-11 repeat was PCR amplified using forward primer #1 [5'-CAG GTA CCA AGG TTT GGC TC-3'] and reverse primer #2 [5'-GTT AAG CTT GTG TTA CAG CTA GAA G-3'] (Table 1). Primer #1 and primer #2 sit 702 bp and 112 bp upstream of the start codon of the first KMP-11 repeat thus generating a 590 bp fragment (A) from the upstream region. The *Kpn*I and *Hind*III restriction digest sites (underlined) that were engineered into primer #1 and primer #2 respectively, facilitated the production of gene knockout cassettes. The PCR was performed using the following conditions: 95° C for 5 minutes (denaturing); followed by 30 cycles of the following: 95° C for 1 minute (denaturing), 55° C for 30 seconds (annealing), 74° C for 30 seconds (extension). The extension was continued at 74° C for 5 minutes prior to holding at 4° C. A second upstream fragment (A') of 698 bp was generated using the reverse primer #3 [5'-GGG AAG CTT GAA AAA GTA TCT T-3'] in conjunction with primer #1. Primer #3, which starts 3 bp upstream of the start site also, contains a *Hind*III restriction site to facilitate cassette cloning. This second PCR fragment encompassed an additional 108 bp of highly conserved DNA that is present upstream of all the KMP-11 gene copies. The second PCR was performed as follows: 95° C for 5 minutes (denaturing); followed by 30 cycles of the following: 95° C for 30 minute (denaturing), 54° C for 30 seconds (annealing), 74° C for 45 seconds (extension), a supplemental extension at 74° C for 5 minutes and a hold at 4° C. Both fragments A and A' were digested with *Kpn*I and *Hind*III (New England Biolabs, Beverly, MA, USA) and cloned into a *Kpn*I/*Hind*III digested pBS neo vector to generate the constructs pBS neo A and pBS neo A', respectively. The next stage was to PCR amplify the KMP-11 3' downstream flanking region from the P1 bacteriophage clone 3D12. The forward 3' downstream primer #7 [5'-GTG GAT CCG AAC ATA TGT ATT TAA G-3'] and reverse primer #8 [5'-CAT CTA GAG CGT CCT CAA CAC G-3'] were designed from sequences 5 bp and 630 bp downstream from the stop codon of the last KMP-11 repeat, respectively. Primer #7 contained a *Bam*HI site while primer #8 contained a *Xba*I site which are underlined. The conditions for the final PCR were as follows: 95° C for 5 minutes (denaturing); followed by 30 cycles of the following: 95° C for 30 seconds (denaturing), 55° C for 30 seconds (annealing), 74° C for 1 minute (extension) and further

extension at 74° C for 5 minutes followed by a hold at 4° C. After isolation, the PCR product termed fragment B was cloned into *Bam*HI/*Xba*I digested pBS neo A and pBS neo A' constructs to generate the completed KMP-11 gene deletion cassettes, pBS neo A/B and pBS neo A'/B. Equivalent disruption cassettes containing the phleomycin drug resistance gene were attained by excision of the neomycin drug resistance gene and insertion of digested (*Hind*III and *Bam*HI) phleomycin drug resistance gene in its place. The phleomycin drug resistance gene used to generate the pBS phleo A/B and pBS phleo A'/B was digested from the previously mentioned pBS phleo construct.

Transfection of trypanosomes via electroporation. Trypanosomes to be used in the transformation were grown for one week prior to electroporation to ensure healthy cultures. The cultures were diluted with an equal volume of fresh medium the day before the electroporation. Trypanosomes (5×10^7) were harvested by centrifugation at 2000xg for 5 minutes at room temperature using a Beckman TJ-R6 tabletop centrifuge (Beckman, Palo Alto, CA, USA), washed once with 5 ml of Zimmerman medium (ZM) to remove traces of FBS, and then resuspended in 500 µl of ZM (0.1 M NaCl, 8mM KCl, 8 mM Na_2HPO_4, 820 µM $Na_2HPO_46H_2O$, 200 µM $Mg(C_2H_3O_2)_24H_2O$, 2.2 µM $Ca(C_2H_3O_2)2H_2O$). The gene deletion cassette to be used (pBS neo A/B, pBS neo A'/B, pBS phleo A/B or pBS phleo A'/B) was excised from the pBluescript vector by a double *Kpn*I and *Xba*I (New England Biolabs, Beverly, MA, USA) restriction digest. Approximately 12 µg of plasmid DNA was digested with 5 units of each enzyme for 4 hours at 37° C. DNA of the appropriate size was isolated from a 1% agarose gel using the Qiaex II Gel Extraction Kit (Qiagen, Chatsworth, CA, USA) yielding about 5µg of specific DNA insert. This was dissolved in 25 µl of sterile distilled water. The DNA and cells were mixed by pipetting and the solution was placed in a 0.2 cm Biorad Gene Pulser/*E.coli* Pulser™ Cuvette (BioRad, Hercules, CA, USA) that had been cooled on ice prior to use. A BioRad Genepulser™ with pulse controller (BioRad, Hercules, CA, USA) was used for the electroporation, pulsing the trypanosomes twice at 1.5 kV, 25 µF, 200 Ω. Deviations from

standard electroporation conditions have been noted (see Results). Between the pulses the cuvettes were vented to remove any potentially volatile gases. Immediately after the second pulse the cuvettes were placed on ice for 5 minutes and subsequently the trypanosomes were transferred to 10 ml of cold procyclic culture medium (containing 20% FBS) and incubated at 27° C.

Selection of stable transfectants. After the trypanosomes recovered from the shock of electroporation (approximately 4 hours), they were centrifuged into a pellet and resuspended into fresh medium containing the appropriate amount of either G418 (a neomycin derivative) or phleomycin. The selection for antibiotic resistant *T.b.brucei* 427.01 started at 10 μg/ml G418 or 5 μg/ml Phleomycin, while for the more sensitive fly transmissible strain *T.b.brucei* 348, 5 μg/ml G418 or 2.5 μg/ml phleomycin was used. Dead trypanosomes were removed from the medium by transferring the culture to a 15 ml conical centrifuge tube and allowing the debris to settle overnight. The selection was then continued on the viable organisms in the supernatant. After 16 to 18 days the drug concentrations were increased to 25 μg/ml G418 or 10 μg/ml phleomycin for *T.b.brucei* 427.01 and 10 μg/ml G418 or 5 μg/ml phleomycin for *T.b.brucei* 348. With healthy growing cultures of transformed *T.b.brucei* 427.01 the antibiotic concentrations were eventually increased to 100 μg/ml G418 or 25 μg/ml phleomycin. The less robust *T.b.brucei* 348 was grown in a final concentration of 50μg/ml G418 or 10 μg/ml phleomycin.

Limiting dilution cloning of transfectants. Once pools of drug resistant trypanosomes were established, individual clones were obtained by dilution cloning. The cells from the pools were counted carefully and then serially diluted (1 in 10) to a concentration of 0.7 trypanosomes/ml in procyclic culture medium containing 20% conditioned medium with the appropriate concentration of the selecting antibiotic. To prepare the conditioned medium, 5

x 10^7 wild type *T.b.brucei* 427.01 were grown in 20 ml of procyclic culture medium in the absence of any antibiotics for 6-8 hours. The culture was centrifuged at 2000x g for 15 minutes at room temperature and the supernatant was removed leaving the trypanosome pellet. The conditioned medium was then sterile filtered through an 0.2 μm Acrodisc® syringe filter (Gelman Sciences, Ann Arbor, MI, USA) to remove any remaining trypanosomes. This conditioned medium was then added to the cloning medium. One hundred microliters of the diluted pool was distributed into each well of two 96 well Falcon 3072 Microtest III™ Tissue Culture Plates (Becton Dickinson Labware, Lincon Park, NJ, USA). The clones were allowed to grow for 7 to 14 days or until substantial growth was visible under the inverted microscope. The entire contents of the microwell cultures were then transferred into 2 ml of the dilution medium in wells of a 24 well Falcon 3072 Microtest III™ Tissue Culture Plate (Becton Dickinson Labware, Lincoln Park, NJ, USA). After several days the clones were diluted into unconditioned medium and were grown up to final culture volumes of 10 ml to 20 ml in 50 ml Falcon tissue culture flasks (Becton Dickinson Labware, Lincoln Park, NJ, USA) before being stored in liquid nitrogen.

Southern blot analysis. For Southern blot analysis, 8 μg of genomic DNA from the *T.brucei* drug resistant transformants were digested overnight at 37°C with 100 units of *Bam*HI (New England Biolabs, Beverly, MA, USA), followed by a further 4 hour incubation at 37°C with an additional 100 units of BamHI to ensure complete digestion. The digests were sodium acetate/ethanol precipitated and electrophoresed for 10 hours at 18 V on a 0.6% agarose gel, along with *Hind*III digested lambda DNA-size standards (New England Biolabs, Beverly, MA, USA). The gel containing the DNA restriction digests was dried down and probed directly. In short, the agarose gel for Southern blot analysis was treated with 1.5 M NaCl/0.5 M NaOH (2 x 30 minutes) and 1.5 M NaCl/0.5 M Tris-HCl pH 7.5 (2 x 15 minutes). The gel was then dried between Whatman filter paper and a piece of Saran wrap using a BioRad model 483 slab dryer (BioRad, Hercules, CA, USA), for 30 minutes without heat and 90 minutes at 80°C. When needed, the gel was rehydrated with

water for 15 minutes, the Saran warp was peeled off and the gel was probed. Approximately 25 ng of the genomic PCR fragment corresponding to the entire KMP-11 coding region (see genomic PCR methods), or of the excised neomycin and phleomycin drug resistance genes (from the pSV2 neo plasmid and the pHD63 vector respectively) were labeled with [α-^{32}P] dATP (specific activity 3000 Ci/mmol) by random priming using the Multiprime DNA labeling system (Amersham, Oakville,ON, Canada) according to the manufacturer's instructions. The Southern blots were treated with prehybridization buffer for 5 minutes at 50°C in 0.25 M Na_2HPO_4/7% SDS and then were incubated overnight at 50°C with mild agitation in the same buffer containing the labeled probe (1.0 x 10^6 cpm/ ml; specific radioactivity 2.0 x 10^9 cpm/μg). The rehydrated gels were washed 2 times for 1 hour each at 50°C with 200 ml of 20 mM Na_2HPO_4 pH 7.2/5% SDS, followed by 2 washes for 20 minutes each at 50°C with 200 ml of 20 mM Na_2HPO_4 pH 7.2/1% SDS. The membranes were left wet, bagged and subsequently autoradiographed using a STORM 820 storage phosphor imaging system (Molecular Dynamics, Sunnyvale, CA, USA).

Generation of external diagnostic primer via dye cycle sequencing. A diagnostic primer outside the KMP-11 flanking regions used in the deletion cassette, was required to aid characterization of drug resistant transformants. A forward sequencing primer #11 [5'-GGA TAT GTT TGT GGA GTA TAA AGT GC-3'] was designed approximately 60 bp from the end of the 3' downstream flanking region and used in dye terminator sequencing. The Dye Terminator Cycle Sequencing Kit with AmpliTaq® DNA Polymerase, FS (Perkin Elmer, Mississauga, ON, Canada) was used in accordance with the manufacturer's instructions. An external diagnostic reverse primer #12 [5'-GGT AGT TTA AGT GAG GGA CTA CG-3'] was designed approximately 50 bp outside of the 3' end of the deletion construct. This primer was used as a definitive test in determining weather the gene knockout cassette in the drug resistance transformants had integrated into the appropriate locus.

Diagnostic PCR testing. Several diagnostic PCR tests were performed on the drug resistant trypanosome transformants. Genomic DNA from the transformants was isolated from trypanosome lysates as described in Chapter 1 (Material and Methods). The PCR amplification mixture (100µl) contained 25pmol each of forward and reverse primers (Gibco BRL, Burlington, ON, Canada), 2mM each of dATP, dCTP, dGTP and dTTP, 1x Taq DNA polymerase buffer, 2.5 units of Taq DNA polymerase (all from Pharmacia, Piscataway, NJ, USA) and 100ng of genomic DNA. Three internal forward primers designed during the production of the KMP-11 gene deletion cassettes and the reverse external diagnostic primer (see previous section) were used to determine the insert location within the trypanosome genome and to test the integrity of the inserted DNA. Forward primers, including primer #6 [5'-CAG GTA CCA AGG TTT GGC TC-3'], designed from the front of the 5' KMP-11 flanking region, primer #13 [5'-CGC AAG CTT ATG ATT GAA CAA GAT GGA-3'] designed from the 5' end of the neomycin drug resistance coding region, and primer #9 [5'-GTG GAT CCG AAC ATA TGT ATT TAA G-3] designed from the 3' downstream flanking region were used in conjunction with the external diagnostic reverse primer #12 [5'-GGT AGT TTA AGT GAG GGA CTA CG-3'] for these PCR tests. The PCR conditions were : 95°C for 5 minutes (denaturing); followed by 30 cycles of the following: 95°C for 1 minute (denaturing), 58°C for 30 seconds (annealing), 74°C for 1 minute (extension). Finally, an extension was continued at 74°C for 5 minutes prior to holding at 4°C. The amplified KMP-11 gene knockout cassette fragments were visualized after agarose gel electrophoresis performed as described in Chapter 1 (Materials and Methods).

Formaldehyde fixation and acetone permeabilization of trypanosomes for immunofluorescence. For formaldehyde fixation, trypanosomes were adjusted to 2 x 10^8/ml in PBS pH 7.4/1% glucose/5% heat-inactivated FBS and were pelleted by

centrifugation at room temperature, 2000x g, for 10 minutes. The supernatants were discarded and the cells resuspended in 10ml of a freshly prepared 4% formaldehyde solution. To generate the 4% formaldehyde solution 4 g of paraformaldehyde was added to 100 ml PBS, heated to 80°C, filtered through Whatman #1 filter paper and cooled to room temperature. The suspension was mixed by inversion and incubated at room temperature for 1 to 2 hours. The fixative was discarded after the trypanosomes were pelleted. Cells were resuspended in 10 ml of PBS. If the cells were to be stored, 0.1% azide was added to the PBS and the cells subsequently underwent immunofluorescence testing as described below. For the acetone permeabilization, trypanosomes were again adjusted to 2 x 10^8/ml in PBS pH 7.4/1% glucose/5% heat-inactivated FBS. Twenty microliters of the cell suspension were added to the slide and then allowed to air dry for 30 minutes. Circles were carved into slides using a diamond pencil to identify the area where the cells were spotted. Slides were placed in pre-chilled acetone (-20^{0}C) and incubated in the freezer for 20 minutes. After the acetone permeabilization the slides were removed from the acetone, air-dried and used for immunofluorescence staining.

Immunofluorescence on fixed/permeabilized trypanosomes. Trypanosomes were adjusted to 2 x 10^8/ml in PBS pH 7.4 /1% glucose / 5% heat-inactivated FBS. If frozen, the first antibody solution was thawed, centrifuged and diluted (neat tissue culture supernatant, 1/500 - 1/1000 ascites fluid or 1/500 immune serum) in PBS / 5% FBS. To examine the KMP-11 gene deletion mutants (*T.b.brucei* neo A/B heterozygotes, and potential homozygous knockouts) for the presence of KMP-11, a mixture of tissue culture supernatants of anti-KMP-11 monoclonal antibodies L98 or L157 were used as the first antibody [Tolson, 1989]. Twenty-five microliters of the first antibody was added to 25ul of the cell suspension in a 12 x 75 mm glass culture tubes and placed on ice. Positive and negative control antibodies as well as controls without first antibody were included in immunofluorescence experiments. The trypanosome/antibody mixture was incubated on ice for 20 minutes and then 1.0 ml of cold PBS / 1% glucose / 5% FBS was added and the

mixture was centrifuged at 2000x g for 10 minutes at 4^0C using a TJ-R6 centrifuge (Beckman, Palo Alto, CA, USA). The supernatants were aspirated off and the washing procedure was repeated once. Fifty microliters of fluorescent second antibody (usually 1/50 - 1/200 dilution of FITC-labeled goat anti-mouse IgG in PBS with 1% glucose and 5% FBS (Caltag, South San Francisco, Ca, USA) were added to the suspension and the tube contents were mixed by vortexing and then incubated on ice for 20 minutes. Three successive washes were performed as previously described, the supernatant was aspirated off and the pellet was resuspended in 20 µl of PBS / 1% glucose / 5% FBS. The cells were spotted on to a slide and allowed to air dry. The SlowFade® Light Antifade Kit (Molecular Probes, Eugene, OR, USA) was employed according to the manufacturer's instructions to reduce the rate at which the fluorescence dye faded under the microscope. Immunofluorescence was observed using a Zeiss Standard binocular microscope fitted with an epifluorescence attachment and a Zeiss Neofluor 63/1.25 oil immersion objective.

RESULTS

KMP-11 gene deletion strategy. The strategy implemented to delete the KMP-11 genes from the trypanosome genome involved the generation of deletion cassettes that could be targeted directly to the KMP-11 locus. The rudimentary steps in the generation of these cassettes are illustrated in figure 19. The isolation and subcloning of *T.brucei* genomic DNA containing the KMP-11 genes from the P1 bacteriophage library has been previously discussed in Chapter 1. Even though four KMP-11 gene repeats was identified within the locus, the tightly packed nature of these genes theoretically should allow their removal in a manner similar to that used to eliminate a single gene. The 5' and 3' flanking regions beyond the first and fourth KMP-11 genes were isolated instead of the DNA surrounding a single gene. This isolation was performed via a PCR amplification of the aforementioned 5' and 3' flanking regions using primer sets with specifically engineered restriction endonuclease sites integrated into the 5' end (Figure 27, page 84; Table 2, page 85). The 5' flanking region of the KMP-11 locus was amplified in a PCR using the forward primer #6 and the reverse primer #7. The 5' UTR forward primer #6 (Figure 27, page 84; Table 2, page 85) contained a *Kpn*I site, which was compatible with the identical site found in the multiple cloning site, directly upstream of the neomycin or phleomycin drug resistance genes in the native pBluescript neomycin and pBluescript phleomycin constructs. The 5' UTR reverse primer #7 (Figure 27, page 84; Table 2, page 85) used in the PCR of the 5' flanking region contained a *Hind* III restriction site that allowed directional insertion of the products into the native pBluescript neomycin and pBluescript phleomycin constructs flush with the start of the drug resistance gene that also contained the *Hind* III site. Similarly, a PCR amplification product of the 3' flanking region was accurately inserted into the target construct just downstream of the drug resistance gene, via the use of a *Bam*HI site and a *Xba*I site (Figure 19, Part 2 & 3). The forward 3' end downstream primer #9 (Figure 27, page 84; Table 2, page 85) containing the *Bam*HI restriction site and the reverse 3' end

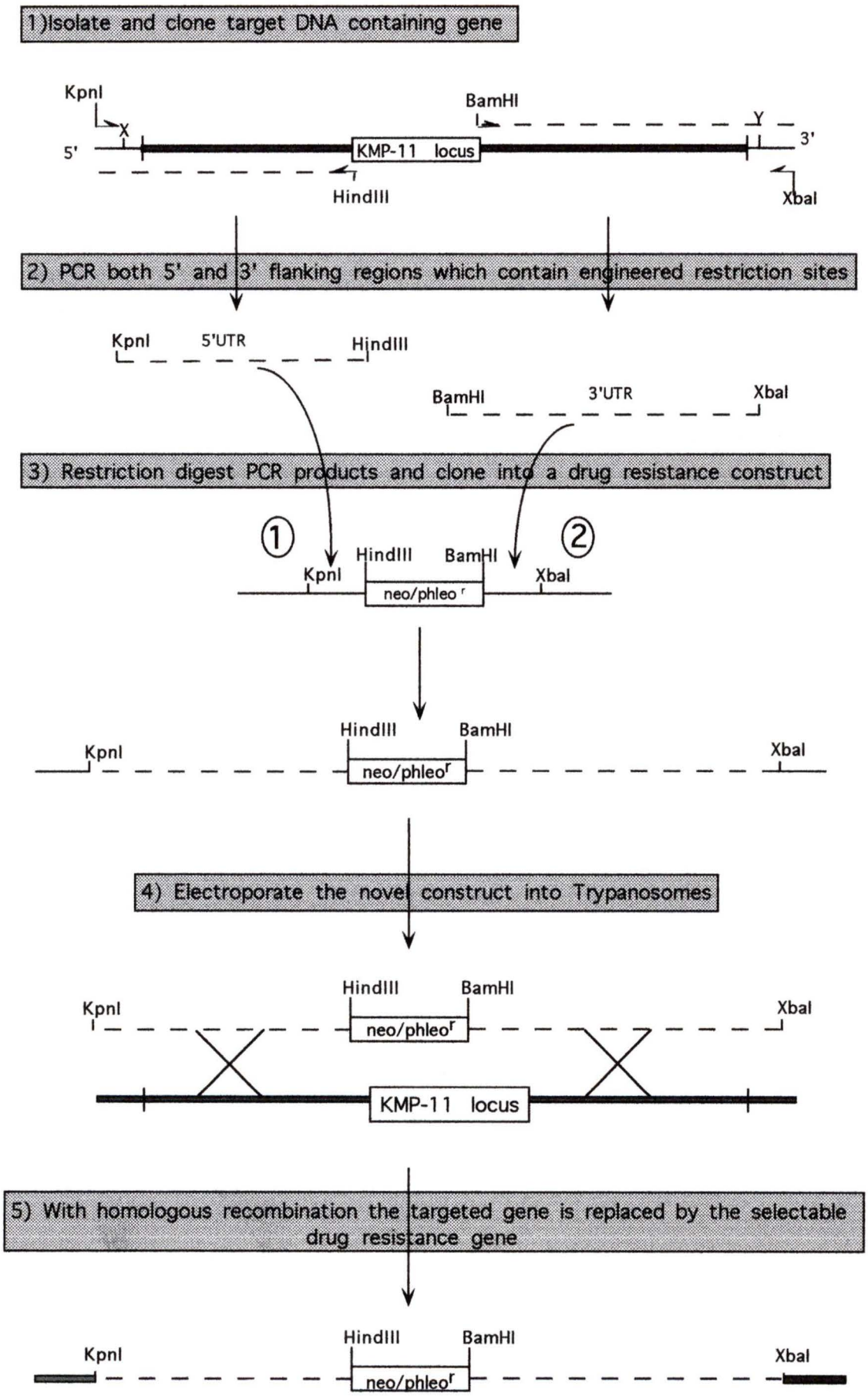


Figure 19: Schematic overview of the methods used to perform the KMP-11 gene knockouts.

downstream primer #10 (Figure 27, page 84; Table 2, page 85) with the *Xba*I restriction site were used to amplify the KMP-11 3' flanking region and integrate it after endonuclease digestion into the precut target construct. The drug resistance construct had corresponding restriction sites: a *Bam*HI site near the stop codon of the drug resistance gene and an *Xba*I site further downstream in the proximal portion of the vector's multiple cloning site (Figure 19, Part 2 & 3). Once the flanking regions of the KMP-11 genes had been inserted on either side of the drug resistance genes the construct was electroporated into the PCF trypanosomes.

Generation of the pBluescript neomycin A/B gene deletion cassette. Plasmids received from Dr. Isabel Roditi's lab in Bern, Switzerland were transformed into *E.coli* XL-1 Blue cells and were subsequently used as stocks of the native drug resistant gene constructs (pBluescript neomycin and pBluescript phleomycin). The 3750 bp pBluescript neomycin construct (Figure 20A, Lane 1) and the 3350 bp pBluescript phleomycin construct (Figure 20A, Lane 2) were linearized to verify the size of the constructs that were to become the backbone of the KMP-11 gene deletion cassettes. PCR amplification of the KMP-11 5' flanking DNA was performed on the pBS KMP-11 3c plasmid using forward primer #6 and reverse primer #7 and produced a 568 bp band (Figure 20, Part B, Lane 2), while the PCR of the 3' flanking region using the forward primer # 8 and reverse primer # 9 produced a slightly larger 605 bp band (figure 20, Part B, Lane 2). The knockout cassette containing the neomycin drug resistance gene was assembled first. The native neomycin construct of ~ 3750 bp (Figure 21, lane 1) was isolated and cleaved in sequential restriction digests using *Kpn*I and *Hind*III endonucleases. The same digestion was performed on the KMP-11 5' flanking region PCR product (Figure 20, Part B, Lane 1), and the insert was ligated into the vector to produce the 4350 bp pBluescript neomycin A construct (Figure 21, lane 2). The new vector was digested with *Bam*HI and then *Xba*I and the precut 3' KMP-11 flanking PCR product (Figure 20, Part B, Lane 2) was directionally cloned into

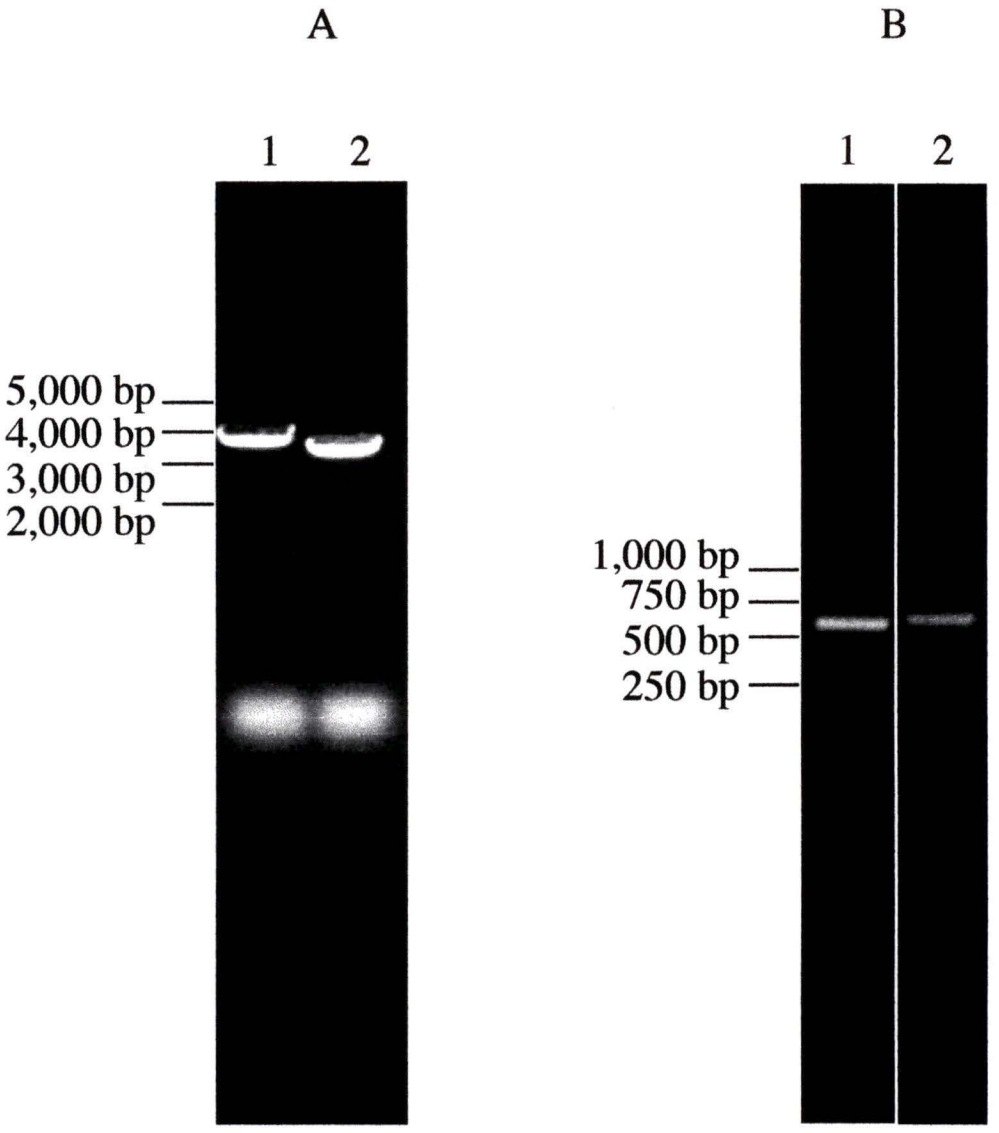
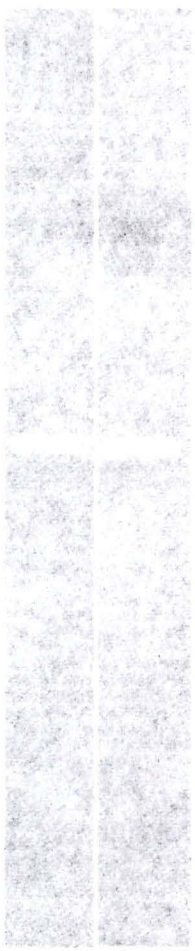


Figure 20: Agarose gel analysis of the pBluescript neomycin (A, Lane 1) and phleomycin (A, Lane 2) constructs linearized using *Hind*III endonuclease. PCR products from the amplification of the KMP-11 5' flanking region (B, Lane 1) and 3' flanking region (B, Lane 2) were also analyzed.



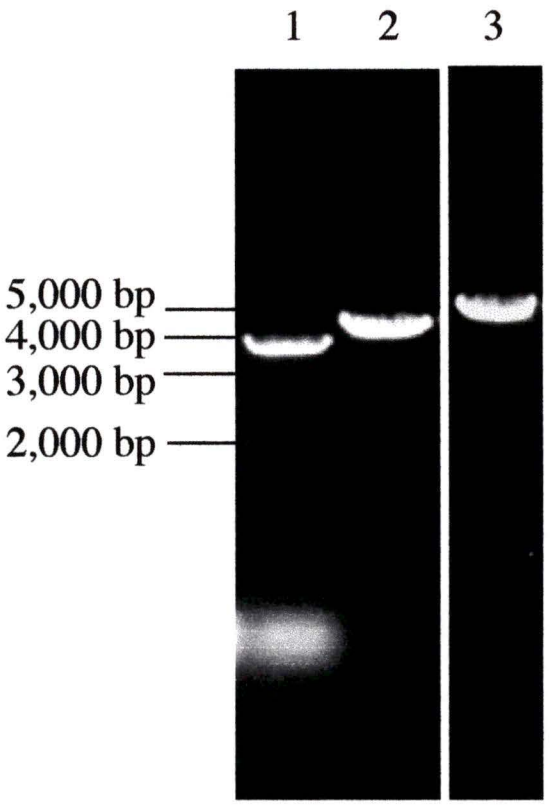


Figure 21: Agarose gel analysis of the drug resistance plasmids. Lane 1, pBluescript neomycin plasmid; Lane 2, pBluescript neomycin A (containing the KMP-11 3' flanking region PCR product); Lane 3, pBluescript neomycin A/B (containing both the KMP-11 3' and 5' flanking region PCR product).

it, producing the final 4950 bp pBluescript neomycin A/B knockout cassette (Figure 21, Lane 3).

Testing of the pBluescript neomycin A/B construct. Restriction digests using both *Kpn*I/*Hind*III and *Bam*HI/*Xba*I endonucleases produced products of approximately 600 bp. Under close scrutiny it is observable that, as expected, the band produced from the latter digest (Figure 22A, Lane 3) was slightly larger than that produced by the former (Figure 22A, lane 1). Three sets of PCR tests were also carried out to characterize the construct. The 3' and 5' flanking regions were amplified using primers #6 and #7 (Figure 22B, Lane 1) and primers #9 and #10 (Figure 22B, Lane 2) respectively, each producing a single band of approximately 600 bp. The entire insert containing the 5' and 3' KMP-11 flanking DNA and the neomycin drug resistance gene were also PCR amplified using the M13 universal forward and reverse primers (Figure 27, Page 84; Table 2, Page 85) producing a band very close to the expected 2000 bp.

Generation of the pBluescript phleomycin disruption cassette. Once the structural integrity of the pBluescript neomycin A/B construct had been verified it was used directly, in a one-stage procedure, to fabricate the pBluescript phleomycin A/B cassette. The completed neomycin construct and the native pBluescript phleomycin construct were doubly digested with the endonucleases *Bam*HI and *Hind*III to remove the respective drug resistance genes. The phleomycin gene band and the cut neomycin construct were isolated and ligated to produce the pBluescript phleomycin A/B gene disruption cassette. The completed constructs were analyzed by agarose gel electrophoresis after being digested with *Bam*HI and *Hind*III (Figure 23). The digestion removed the 798 bp neomycin (Figure 23, Lane 1) and 375 bp phleomycin (Figure 23, Lane 2) drug resistance genes leaving the bulk (~4000 bp) of the construct intact. Thus, both a neomycin and a phleomycin disruption cassette had been generated and contained trypanosome genomic DNA from regions flanking the

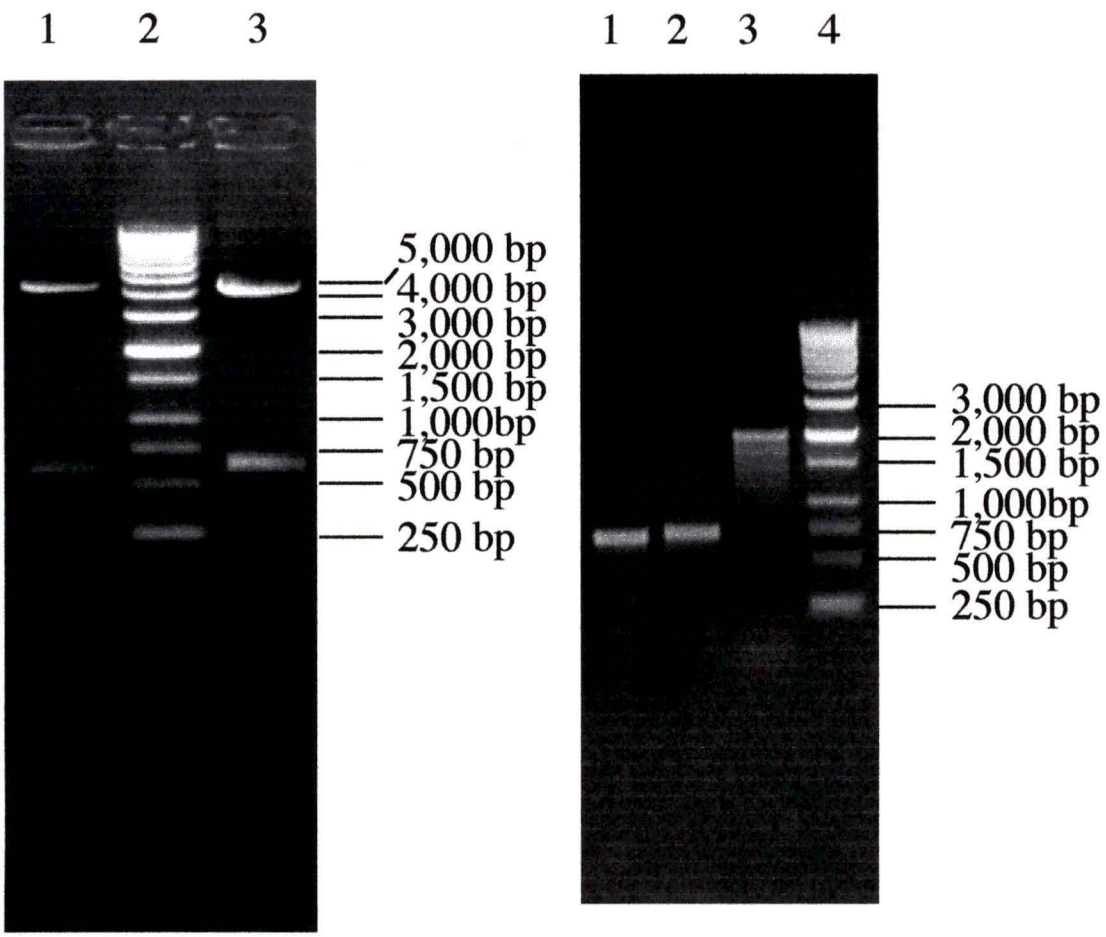


Figure 22: Agarose gel analysis of the pBluescript neomycin A/B construct after restriction digestion or after PCR analysis. A, Restriction digests. Lane 1, *Kpn*I and *Hind*III; Lane 2, kb standards; Lane 3, *Bam*HI and *Xba*I digestion. B, PCR analysis. Lane 1, primers #6 and#7; Lane 2, primers #9 and #10; Lane 3, primers #4 and #5; Lane 4, kb standards. The sizes of the kb DNA ladder standards are indicated to the right of the panels.

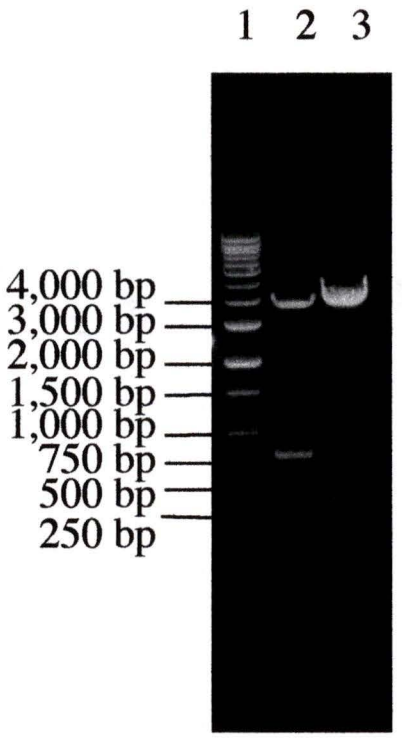


Figure 23: Agarose gel analysis of final pBluescript neomycin A/B and pBluescript phleomycin A/B constructs. Both constructs were digested with *Hind* III and *Bam*HI endonucleases to remove the neomycin or phleomycin drug resistance genes respectively. Lane 1, kb standards; Lane 2, pBluescript neomycin A/B; Lane 3, pBluescript phleomycin A/B. The bp standards are shown to the left of the panel.

KMP-11 locus. Each drug resistance gene is flanked by 586 bp of upstream KMP-11 locus flanking sequence and 605 bp of downstream flanking sequence (see Figure 26, Page 83).

Generation of a second set of drug resistance disruption cassettes. In much the same way that the pBluescript phleomycin A/B construct was generated by removing the neomycin drug resistance gene and replacing it with the phleomycin resistance gene, a new set of constructs containing an extended 5' KMP-11 flanking region was generated. The original 5' upstream flanking PCR product was replaced with an extended DNA fragment which contained an extra region of nucleic acid sequence from immediately upstream of the KMP-11 gene start codon. To generate this new fragment a 5' end upstream reverse primer #8 (Figure 27, page 84; Table 2, page 85) was designed and used in concert with the forward primer #6. The novel fragment thus contained the conserved upstream untranslated region of 102 bp that had been previously identified (See Chapter 1) as the repeat in front of each KMP-11 gene. Since primer #8 was designed from a sequence that appeared several times in the KMP-11 locus the PCR amplification of the novel fragment produced several products (Figure 24A, Lane 1) of which the desired fragment was the smallest and most intense band of around 700 bp. The lowest molecular weight PCR product once isolated (Figure 24B, Lane 1) was shown to be slightly larger than the 568 bp fragment (Figure 24B, Lane 2) that it was replacing. The pBluescript neomycin A/B and pBluescript phleomycin A/B constructs were *Kpn*I/*Hind*III digested (Figure 25A, Lane 1 & 2) to remove the original 5' flanking region and the similarly digested novel PCR product was cloned directionally into the *Hind*III-*Kpn*I cut site. The new constructs termed pBluescript neomycin A'/B and pBluescript phleomycin A'/B contained DNA homologous to sequences in several locations in the KMP-11 locus and thus may integrate differently to the original constructs. As expected, when digested with *Kpn*I and *Xba*I to remove the entire insert, pBluescript neomycin A'/B produced a band of greater

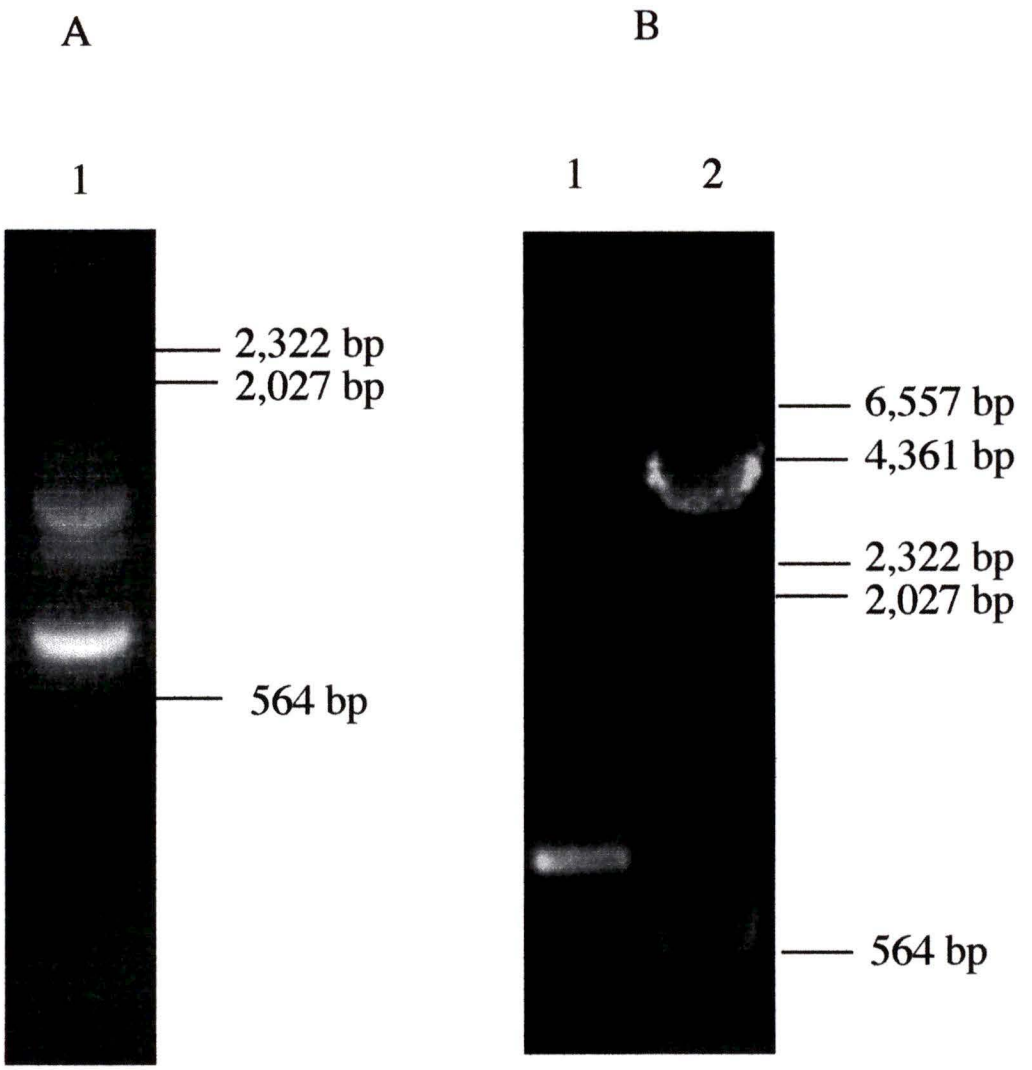
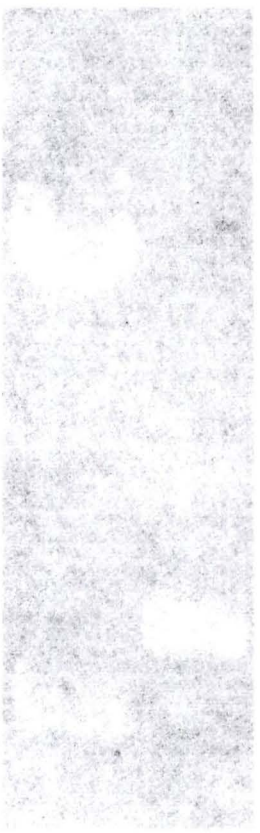


Figure 24: Panel A shows an agarose gel analysis of the products from a PCR amplification of the 5' KMP-11 flanking region. Panel B shows the 710 bp KMP-11 5' flanking region PCR product (B, Lane 1) which replaced the initial 568 bp KMP-11 5' flanking region (B, Lane 2) from the pBluescript neomycin A/B or the pBluescript phleomycin A/B construct when the new set of pBluescript neomycin A'/B construct and pBluescript phleomycin A'/B constructs were generated.



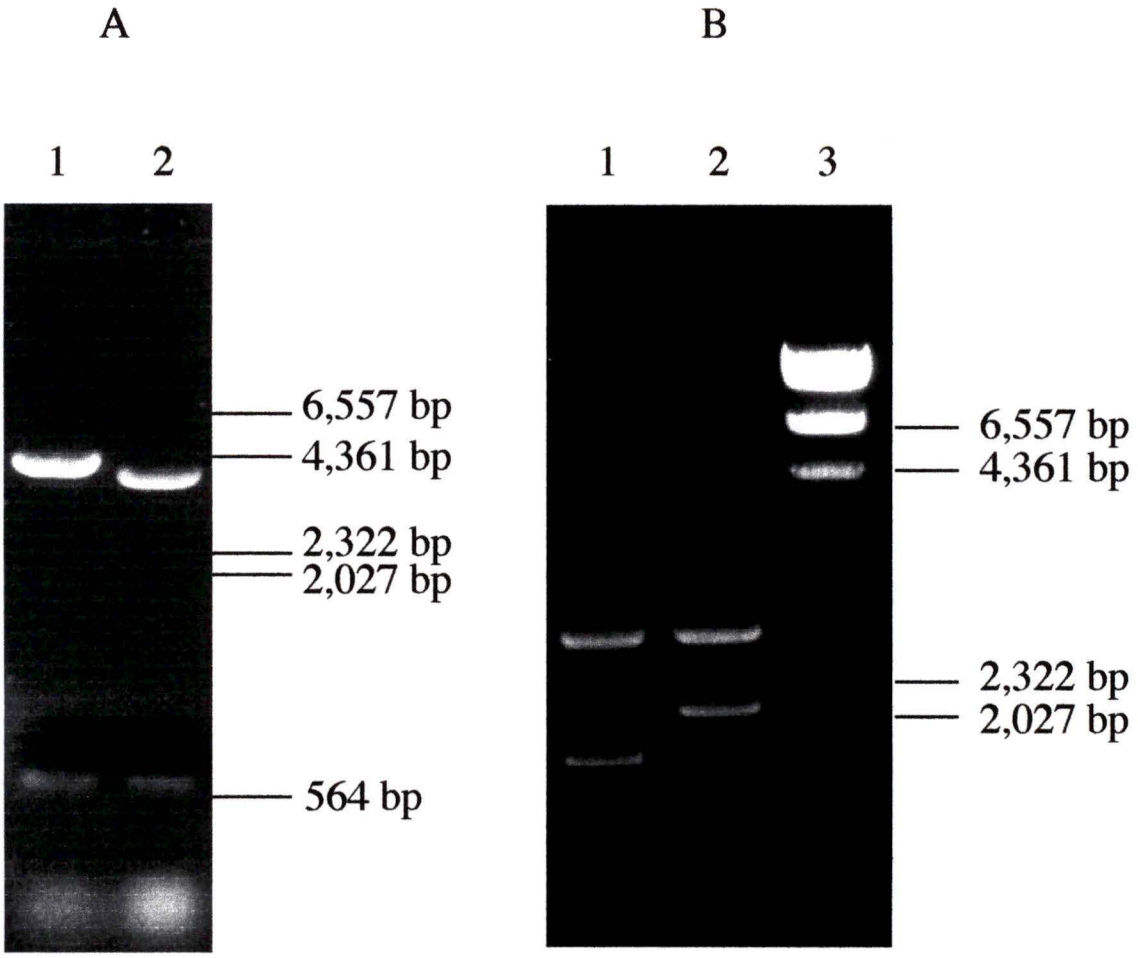


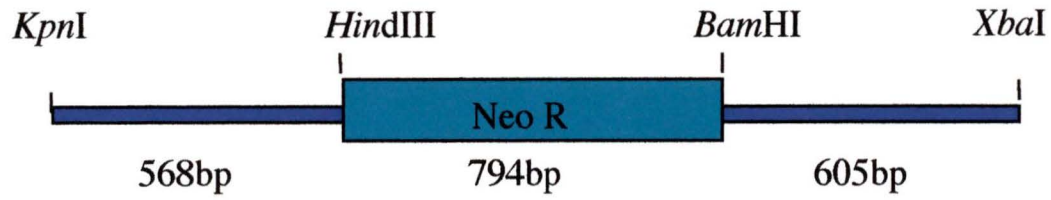
Figure 25: Agarose gel analysis of the pBluescript neomycin drug resistance constructs. A, Lane 1, *Kpn*I/*Hind*III digested pBluescript neomycin A/B construct. A, Lane 2, *Kpn*I / *Hind*III digested pBluescript phleomycin A/B construct. These digested constructs were subsequently used to generate of the "A'/B" constructs. B, Lane 1, *Kpn*I/*Xba*I digested pBluescript phleomycin A'/B construct. B, Lane 2, *Kpn*I/*Xba*I digested pBluescript neomycin A'/B construct. B, Lane 3, Molecular size standards are indicated on the right of the panels.

than 2000 bp (Figure 25B, Lane 1) while the pBluescript phleomycin A'/B insert produced was approximately 1700 bp (Figure 25B, Lane 2).

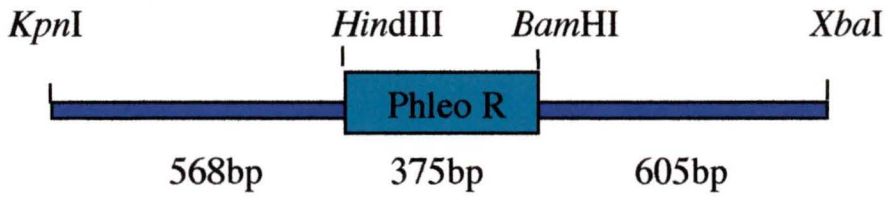
In total, four KMP-11 gene deletion cassettes were generated. The cassettes contained a selectable drug resistance gene marker (neomycin or phleomycin) surrounded by flanking regions from the KMP-11 locus (Figure 26). The production of these constructs was the groundwork required before attempting gene knockout studies of the KMP-11 locus in the African trypanosome.

Electroporation and selection of wild type trypanosomes using the neomycin drug resistance gene constructs. Approximately 5 ug of insert from either the pBluescript neomycin A/B or the pBluescript neomycin A'/B constructs were isolated and electroporated into 5 X 10^7 *T.b.brucei* 427.01 PCF trypanosomes. The selection procedure was started four hours after the transformation when 25 ug/ml of G418 (a potent neomycin derivative) was added to the trypanosome culture medium. For the first week the culture medium was "refreshed" daily as the cultures became continuously more dense and the medium became acidic as indicated by its yellow color. Refreshing the culture was simply performed by centrifuging the trypanosomes and resuspending them in fresh medium containing 25 ug/ml of G418. On the 6th day of the neomycin (G418) selection the trypanosomes started to decline in numbers. The debris caused by the necrosis of the trypanosomes was removed periodically thereafter by placing the culture in a 50 ml conical tube overnight and retaining the top nine-tenths of the supernatant while removing the pellet. By the ninth day post electroporation, very few healthy looking, rapidly dividing trypanosomes were visible in the culture. The dead and dying trypanosomes had drastically altered morphology, becoming either very thin and elongated (approximately twice the normal length) or rounded up having a spherical central body cavity with protuberances on either side. Substantial regrowth of the few surviving drug resistant trypanosomes was evident by the 13th day from both the pBluescript neomycin A /B and A'/B transformations. The pools of drug resistant trypanosomes were subject to two

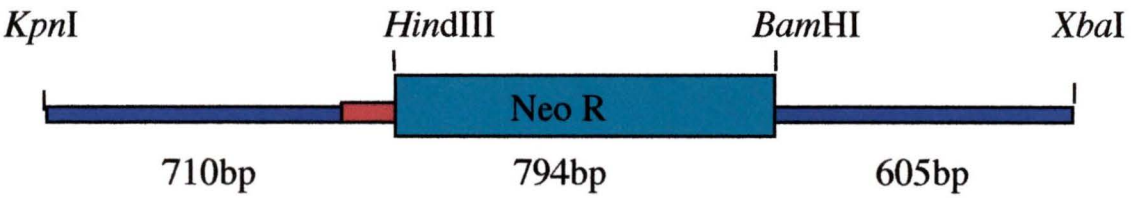
a) pBluescript neo A/B construct



b) pBluescript phleo A/B construct



c) pBluescript neo A'/B construct



d) pBluescript phleo A'/B construct

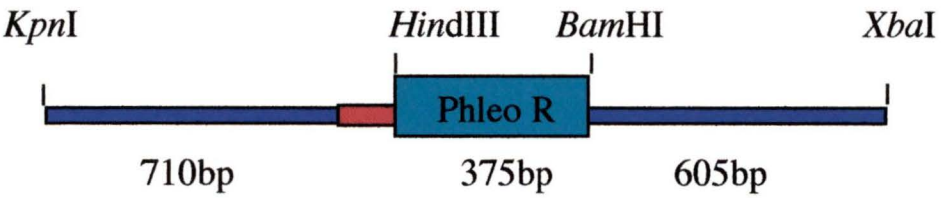


Figure 26: Schematic representation of the pBluescript neomycin and pBluescript phleomycin constructs used in gene deletion experiments.

- Neomycin or Phleomycin drug resistance gene
- KMP-11 gene locus 3' and 5' flanking genomic DNA
- Conserved untranslated genomic region found upstream of each KMP-11 gene

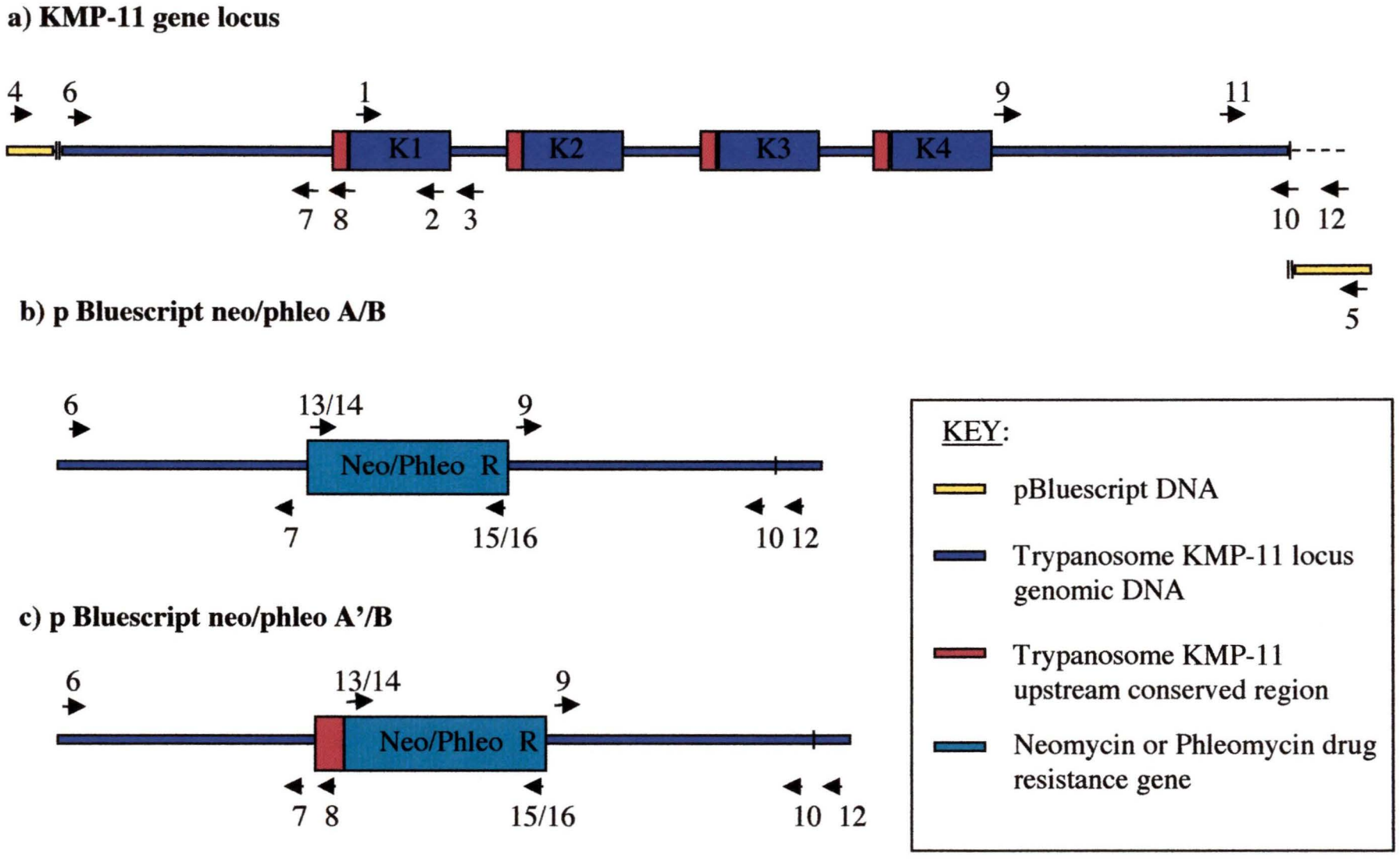


Figure 27: Schematic of primers generated for KMP-11 locus sequencing and gene deletion studies.

	Primer Name	Primer use	Primer Sequence
Primer #1	KMP-11 gene amplification forward	Southerns, Diagnostic PCR	5'-ATG GCC ACC ACA TAC GAA G-3'
Primer #2	KMP-11 gene amplification reverse A	Southerns, Diagnostic PCR	5'-TCA TTT TCC GGG GAA CTG-3'
Primer #3	KMP-11 gene amplification reverse B	Southerns, Diagnostic PCR	5'-AAT GGA AAG AAA ATG AGA GGT G-3'
Primer #4	pB.S. forward universal primer (-20)	PCR, Shot gun Sequencing	5'-GTA AAA CGA CGG CCA GT-3'
Primer #5	pB.S. reverse universal primer (M13)	PCR, Shot gun Sequencing	5'-CAG GAA ACA GCT ATG AC-3'
Primer #6	5' end upstream forward	Gene deletion construct, diagnostic PCR	5'-CAG GTA CCA AGG TTT GGC TC-3'
Primer #7	5' end upstream reverse (A)	Gene deletion construct	5'-GTT AAG CTT GTG TTA CAG CTA GAA G-3'
Primer #8	5' end upstream reverse (B)	Gene deletion construct	5'-GGG AAG CTT GAA AAA GTA TCT T-3'
Primer #9	3' end downstream forward	Gene deletion construct diagnostic PCR	5'-GTG GAT CCG AAC ATA TGT ATT TAA G-3'
Primer #10	3' end downstream reverse	Gene deletion construct	5'-CAT CTA GAG CGT CCT CAA CAC G-3'
Primer #11	3' end sequencing forward	Sequencing of 3' end external DNA	5'-GGA TAT GTT TGT GGA GTA TAA AGT GC-3'
Primer #12	3' end external diagnostic reverse	Diagnostic PCR	5'-GGT AGT TTA AGT GAG GGA CTA CG-3'
Primer #13	Neomycin resistance gene amp. forward	Sotherns, Diagnostic PCR	5'-CGC AAG CTT ATG ATT GAA CAA GAT GGA-3'
Primer #14	Phleomycin resistance gene amp. forward	Sotherns, Diagnostic PCR	5'-GGG ATG GCC AAG TTG ACC A-3'
Primer #15	Neomycin resistance gene amp. reverse	Sotherns, Diagnostic PCR	5'-TAA GGA TCC TCA GAA GAA CTC GTC-3'
Primer #16	Phleomycin resistance gene amp. reverse	Sotherns, Diagnostic PCR	5'-TCA GTC CTG CTC CTC GGC C-3'

Table 2: Primers used during the KMP-11 characterization and gene deletion experiments.

successive drug concentration increases, to 50 ug/ml and then to 100 ug/ml at which concentration they were maintained.

Isolation of single neomycin resistant transformants. Dilution cloning of the drug resistant pools of transformants that had been electroporated with the pBluescript neomycin A/B and pBluescript neomycin A'/B constructs produced 5 and 11 individual clones respectively. Several of the fast growing transformants including *T.b.brucei* neo A/B clone 1 and clone 4 as well as *T.b.brucei* neo A'/B clones 1, 2, 10 and 11 were chosen for further analysis.

Genetic characterization of the neomycin resistant trypanosome clones. Southern blot and genomic PCR analysis were performed to characterize the type of integration event that had taken place and the location of the insert within the genome. For Southern blot analysis, the transformant genomic DNA was isolated, digested with *Bam*HI, electroporated on a 0.6% agarose gel and the gel was dried before being probed. The first probe used was the 279 bp complete KMP-11 gene fragment. The *Bam*HI digested *T.b.brucei* neo A/B clone 1 (Figure 28, Lane 1) and clone 4 (Figure 28, Lane 2) genomic DNA shows a single band of 3500 bp hybridized to the KMP-11 probe. The same 3500 bp band is visible for the untransformed wild type *T.b.brucei* 427.01 trypanosomes (Figure 28, Lane 7) and represents the native 3375 bp KMP-11 locus that is flanked by two *Bam*HI sites in the genomic DNA (Figure 15 and 16). The *T.b.brucei* neo A'/B clones 1, 2, 10, and 11 (Figure 28, Lanes 3 - 6) all show a banding pattern that is similar to one another, with the previously mentioned 3500 bp band as well as a single band of lower intensity of about 2100 bp. A second Southern blot (Figure 29) was then performed using the neomycin resistance gene as a probe. The *T.b.brucei* neo A/B clones 1 and 4 (Figure 29, Lanes 1 and 2) show a single band of approximately 1400 bp hybridized to the neomycin drug resistance gene probe. The two *T.b.brucei* neo A'/B clones tested, clone 1 and clone 2 (Figure 29, Lanes 3 and 4), also showed a single band. Interestingly, this band is the same size (2100 bp) as the lower molecular weight band in the KMP-11 probed Southern blot (Figure 28, Lanes 3 and 4). The wild type trypanosome control (Figure 29, Lane 5), as

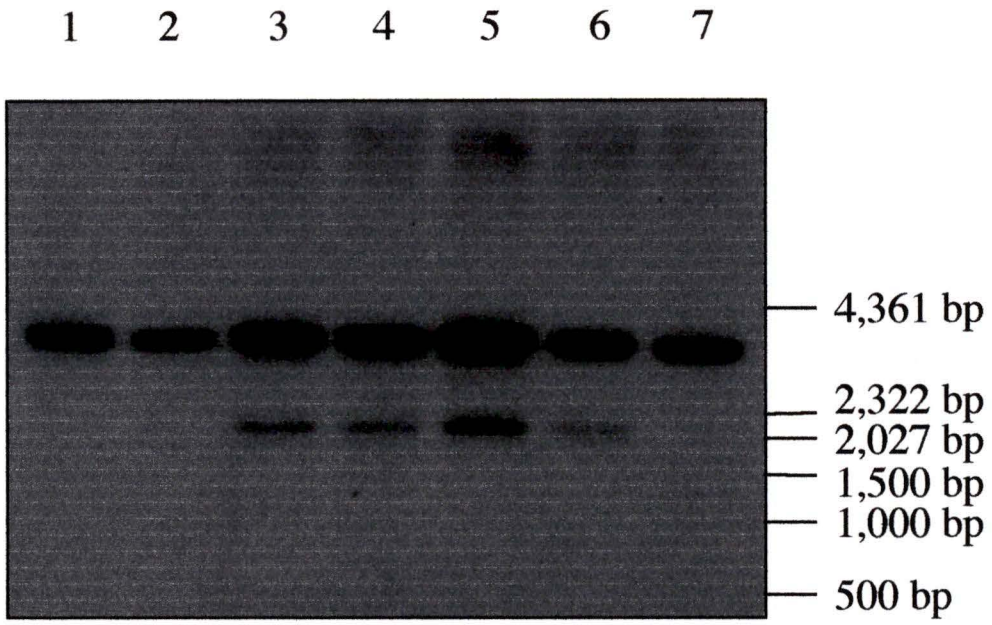


Figure 28: Southern blot analysis of *T.b.brucei* 427 DNA after the first round of KMP-11 gene knockouts. Trypanosome genomic DNA was digested with BamHI, electrophoresed on a 0.6% agarose gel and probed with the complete ^{32}P-labeled KMP-11 gene. The Southern blot contains genomic DNA from trypanosomes that were electroporated with the pBS neo A/B vector (lanes 1-2), with the pBS-neo A'/B vector (lanes 3-6), or without any vector (lane 7).
Size standards are shown on the right of the figure.

Lane 1: *T.b.brucei* neo A/B clone 1, *T.b.brucei* 427.01 electroporated with pBS neo A/B

Lane 2: *T.b.brucei* neo A/B clone 4, *T.b.brucei* 427.01 electroporated with pBS neo A/B

Lane 3: *T.b.brucei* neo A'/B clone 1, *T.b.brucei* 427.01 electroporated with pBS neo A'/B

Lane 4: *T.b.brucei* neo A'/B clone 2, *T.b.brucei* 427.01 electroporated with pBS neo A'/B

Lane 5: *T.b.brucei* neo A'/B clone 10, *T.b.brucei* 427.01 electroporated with pBS neo A'/B

Lane 6: *T.b.brucei* neo A'/B clone 11, *T.b.brucei* 427.01 electroporated with pBS neo A'/B

Lane 7: *T.b.brucei* 427.01 non-electroporated control

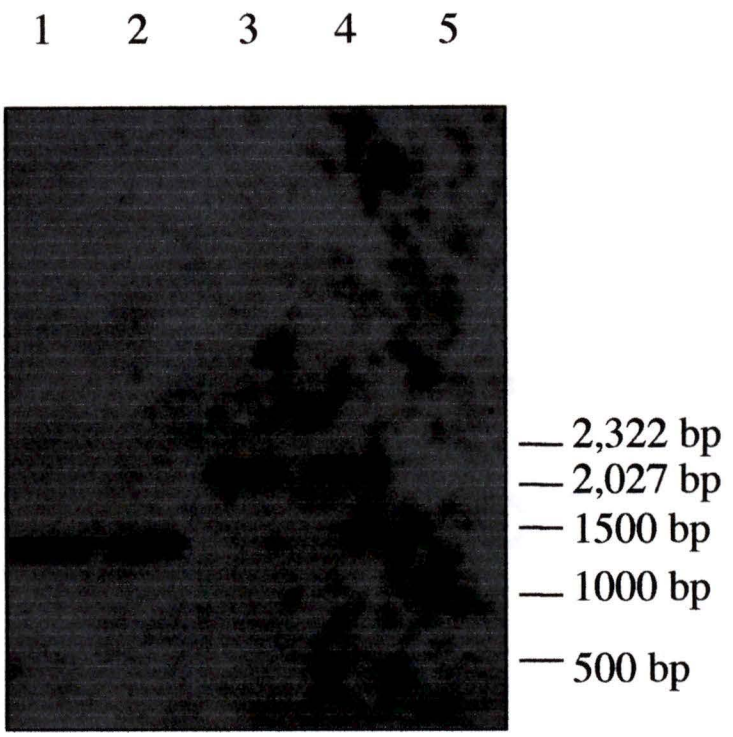


Figure 29: Southern blot analysis of transformed trypanosomes, using the neomycin resistance gene as a probe. Trypanosome genomic DNA was digested with BamHI, electrophoresed on a 0.6% agarose gel and probed with the complete ^{32}P-labeled KMP-11 gene. Lanes 1 through 4 contain DNA from the same trypanosome clones as in the previous figure and lane 5 contains the wild type control. Size standards are shown on the right of the figure.

Lane 1: *T.b.brucei* neo A'/B clone 1, *T.b.brucei* 427.01 electroporated with pBS neo A'/B

Lane 2: *T.b.brucei* neo A'/B clone 4, *T.b.brucei* 427.01 electroporated with pBS neo A'/B

Lane 3: *T.b.brucei* neo A/B clone 1, *T.b.brucei* 427.01 electroporated with pBS neo A/B

Lane 4: *T.b.brucei* neo A/B clone 2, *T.b.brucei* 427.01 electroporated with pBS neo A/B

Lane 5: *T.b.brucei* 427.01 non-electroporated control

expected, showed no bands that hybridized with the neomycin resistance gene probe. Taken together the information from the two Southern blots tells an interesting story of how the different deletion cassettes integrated into the trypanosome genome: The *T.b.brucei* neo A/B transformants have likely had the neomycin A/B construct integrate into the KMP-11 locus generating a complete knockout mutant, that has had one of its KMP-11 gene loci deleted. First, the lack of any secondary bands on the KMP-11 probed Southern indicates that no partial disruption event took place. Second, the size (1400 bp) of the band hybridizing to the neomycin resistance gene probe is exactly as expected for an accurate integration of the deletion cassette into the KMP-11 locus (Figure 30A). If this was the case, the transformant DNA run on the Southern would have been cut once within the insert at the *Bam*HI site near the end of the neomycin resistance gene and a second time at the genomic DNA *Bam*HI site about 50 bp upstream of the integration site. This scenario (Figure 30A) could have produced the 1420 bp hybridizing band that was evident on the neomycin probed Southern blot (Figure 29, Lanes 1 and 2).

It is evident from the Southern blotting data that the *T.b.brucei* neo A'/B transformant underwent a completely different homologous recombination event than the *T.b.brucei* neo A/B transformants which produced a partial deletion of the targeted KMP-11 locus. Both Southern blots showed a hybridizing band of approximately 2100 bp indicating that the KMP-11 and the neomycin resistance gene were present on the same *Bam*HI generated fragment. From the fragment size it is apparent that the A'/B deletion cassette integrated into the KMP-11 locus leaving a lone KMP-11 gene intact (Figure 30B). Again, the DNA was cut at the neomycin resistance gene *Bam*HI site inside the deletion cassette and at external genomic *Bam*HI site, but because of the added KMP-11 gene and flanking sequence, the size of the fragment increased to 2100 bp. The repeated upstream untranslated region, that was added to the pBluescript neomycin A'/B construct which is found in front of each KMP-11 gene repeat, may have facilitated the partial locus deletion.

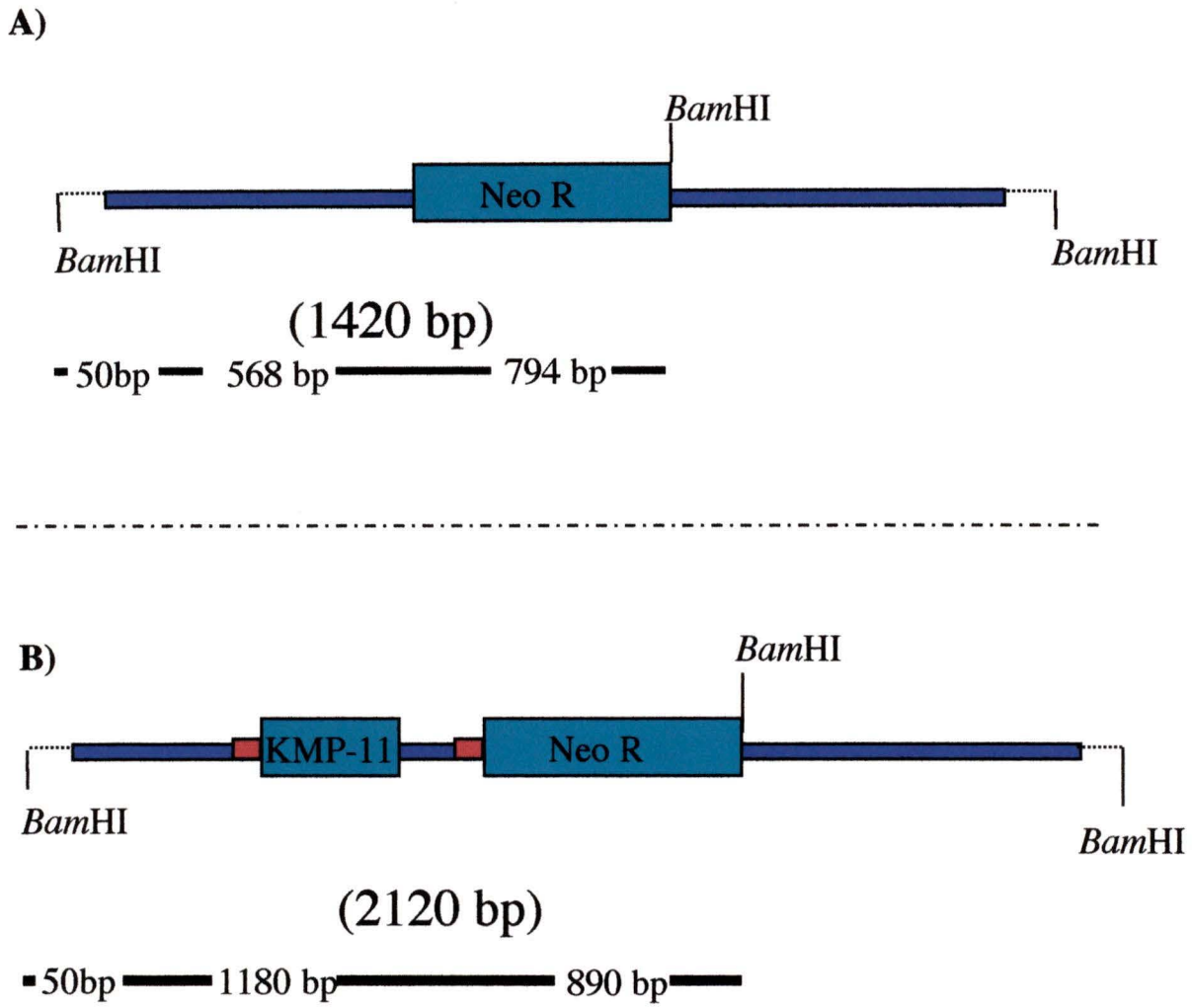


Figure 30: Interpretation of the Southern blot results from the first round of KMP-11 locus targeted gene deletions using neomycin as the selectable marker. The neomycin drug resistance gene was used to probe the blot. The wild type *T.b.brucei* 427.01 PCF were electroporated with (A) pBS neo A/B generating full heterozygote knockouts or (B) pBS-neo A'/B which lead to a partial knockout heterozygote that still contained one KMP-11 gene in the targeted locus.

- Neomycin or Phleomycin drug resistance gene
- KMP-11 gene locus 3' and 5' flanking genomic DNA
- Conserved untranslated genomic region found upstream of each KMP-11 gene
- PCR amplification products

PCR testing of the neomycin A/B transformants to localize the deletion cassette to the KMP-11 locus. A diagnostic PCR test was performed on the *T.b.brucei* neo A/B clones using forward primers internal to the insert and an external reverse diagnostic primer (Figure 27, Table 2) that had been designed from genomic DNA downstream from the KMP-11 locus. Three PCRs were performed using the diagnostic reverse primer #12. Forward primers #6, #13 and #9 were used with the reverse diagnostic primer #12 in three separate PCR reactions on *T.b.brucei* neo A/B clones 1 and 2 genomic DNA and on the wild type *T.b.brucei* 427.01 PCF control DNA. The PCRs of *T.b.brucei* neo clone 1 and clone 2 using the forward primer #6 and the external reverse primer #12 generated a product of 2000 bp in each reaction (Figure 31, Lane 1 and 2). The PCR using the forward primer #13 and reverse primer #12 showed products of approximately 1400 bp (Figure 31, Lane 4 and 5). The same PCRs using wild type genomic DNA produced no visible products (Figure 31, Lane 3 and 6). The PCR using the forward primer #9 and reverse primer #12 produced a 700 bp band for the *T.b.brucei* neo A/B clone 1(Figure 31, Lane 8) but none for clone 2 (Figure 31, Lane 9) or the wild type control (Figure 31, Lane 10). The absence of the 700 bp band from clone 2 in this last reaction was likely due to an improperly prepared reaction mixture or lack of Taq polymerase addition. Apart from this last exception, the bands produced in the PCR testing were consistent with a complete and accurate integration of the deletion cassette into the KMP-11 locus (Figure 32).

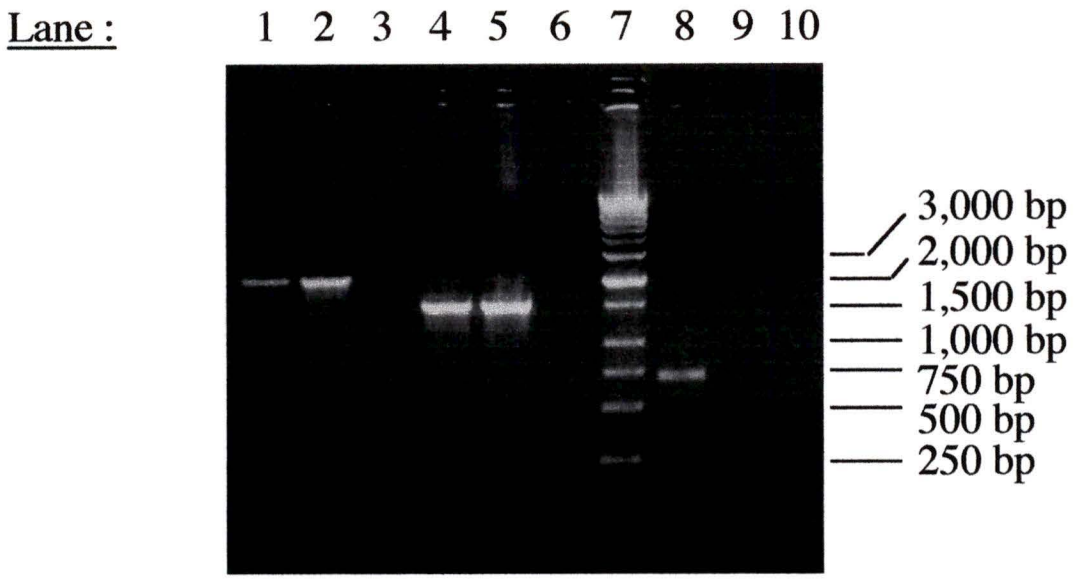


Figure 31: Agarose gel electrophoresis of products from diagnostic PCRs of neomycin transformants. Lane 1, 4, 8: *T.b.brucei* neo A/B clone 1 Lane 2, 5, 9:*T.b.brucei* neo A/B clone 4 and Lanes 3, 6, 10: *T.b.brucei* 427 wild type non-transformed trypanosomes. Lanes 1 to 3 contain PCR products from reactions that used primer #6 (KMP-11 5' upstream forward) and primer #12 (external diagnostic reverse). Lanes 4 to 6 contain PCR products from reactions using primer #13 (neomycin resistance gene amp. forward) and primer #12 (external diagnostic reverse). Lanes 8 to10 contain PCR products from reactions using primer #9 (KMP-11 3' downstream reverse) and primer #12 (external diagnostic reverse). Lane 7 contained Kb DNA ladder size standards. The location of the primers is schematically illustrated in Figure 27.

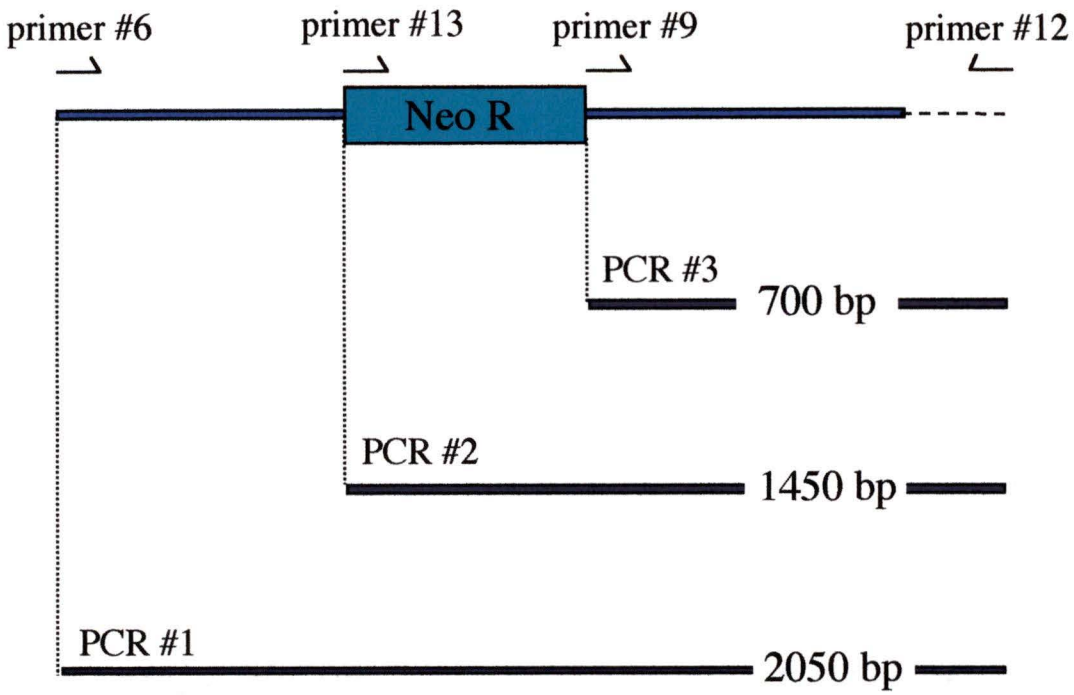


Figure 32: Interpretation of the results from the diagnostic PCR (see previous figure, 31). The PCRs were performed on the neomycin drug resistant transformants *T.b.brucei* neo A/B clone 1 and *T.b.brucei* neo A/B clone 2 as previously described. The three sets of PCR that were performed on the transformants produced bands of sizes 700bp, 1450bp, and 2050bp which was consistent with a complete deletion of the KMP-11 gene locus.

Neomycin drug resistance gene

KMP-11 gene locus 3' and 5' flanking genomic DNA

PCR amplification products

Second round of KMP-11 locus deletions using the phleomycin drug resistance constructs. For the second round of knockouts, *T.b.brucei* neo A/B clone 1 and *T.b.brucei* neo A'/B clone 2 were electroporated with either pBluescript phleomycin A/B or pBluescript phleomycin A'/B. An initial set of electroporations using the neomycin transformants was unsuccessful, thus subsequent electroporation and selection conditions were varied to achieve the desired results. Five μg to 20 μg of insert from the pBluescript phleomycin A/B or the pBluescript phleomycin A'/B construct were electroporated into either *T.b.brucei* neo A/B clone 1 or *T.b.brucei* neo A'/B clone 2. Electroporation included two pulses at one of the following 1.5 kV, 1.8 kV, 2.0 kV or 2.5 kV, or three pulses at either 1.0 kV or 1.5 kV. The selection procedure was started 4 hours after electroporating by adding 100 μg/ml of G418 and 2.5 μg/ml phleomycin to the growth medium. The double drug transformants were grown in MEM plus 15 % FBS and 20 % conditioned medium, providing a rich environment for growth of the compromised trypanosomes. For the first 5 to 7 days, the rapidly growing cultures of trypanosomes required daily "refreshing" of the media, after which time the numbers of trypanosomes dwindled as the drug took effect. The regrowth of the surviving drug resistant trypanosomes was unique for each electroporation. *T.b.brucei* neo A'/B clone 2 cultures electroporated with either the pBluescript phleomycin A/B insert or the pBluescript phleomycin A'/B insert showed substantial regrowth within 18 days. In contrast, most *T.b.brucei* neo A/B clone 1 cultures electroporated with the pBluescript phleomycin A/B insert were viable but showed retarded regrowth over the same period.

Several stable transformant pools have been attained and are undergoing further characterization, while others are still in the process of regrowing after the drug selection (Table 3). A pool of *T.b.brucei* neo A'/B transformants electroporated with the insert from the pBluescript phleomycin A'/B construct (Pool G) was the first to regrow to a regular population density, two weeks post electroporation. Within several days, other transformant pools showed similar growth. These included *T.b.brucei* neo A'/B transformants electroporated with 5μg (Pool J) or 10μg (Pool K) of pBluescript phleomycin A/B insert as well as *T.b.brucei* neo A/B transformants electroporated with

	Transformants	Deletion Construct (pBluescript)	Amount of DNA (μg)	Number of Pulses & Voltage (kV)	Growth of Pool
Pool A	*T.b.brucei* neo A/B clone 1	phleomycin A/B	5	2 X 2.0	+
Pool B	*T.b.brucei* neo A/B clone 1	phleomycin A/B	10	2 X 1.8	~
Pool C	*T.b.brucei* neo A/B clone 1	phleomycin A/B	20	2 X 1.5	~
Pool D	*T.b.brucei* neo A/B clone 1	phleomycin A'/B	5	2 X 1.8	+
Pool E	*T.b.brucei* neo A/B clone 1	phleomycin A/B	10	2 X 1.8	+
Pool F	*T.b.brucei* neo A'/B clone 2	phleomycin A/B	5	3 X 1.0	~
Pool G	*T.b.brucei* neo A'/B clone 2	phleomycin A'/B	5	2 X 1.8	+
Pool H	*T.b.brucei* neo A/B clone 1	phleomycin A/B	5	2 X 1.5	~
Pool I	*T.b.brucei* neo A/B clone 1	phleomycin A/B	10	2 X 1.5	~
Pool J	*T.b.brucei* neo A'/B clone 2	phleomycin A/B	5	2 X 1.5	+
Pool K	*T.b.brucei* neo A'/B clone 2	phleomycin A/B	10	2 X 1.5	+
Pool L	*T.b.brucei* neo A/B clone 1	phleomycin A/B	5	3 X 1.5	+
Pool M	*T.b.brucei* neo A/B clone 1	phleomycin A/B	10	2 X 1.5	~
Pool N	*T.b.brucei* neo A/B clone 1	phleomycin A/B	5	2 X 2.0	~
Pool O	*T.b.brucei* neo A/B clone 1	phleomycin A/B	10	2 X 2.5	~
Control	*T.b.brucei* neo A/B clone 1	None	-	2 X 1.5	-

Table 3: Tabulation of the transformations performed for the second round of KMP-11 gene locus knockouts. The growth of the transformant pools were rated as follows: "+" indicates that the transformants attained densities associated with normal healthy cultures, "~" indicates that the cultures are viable but struggling after 28 days and "-" indicates that the culture died within 16.

pBluescript phleomycin A'/B insert (Pool D). Electroporations of *T.b.brucei* neo A/B transformants using varying amounts of pBluescript phleomycin A/B insert produced dense populations of transformants (Pools A, E, O) showing strong growth within 21 days. Still other electroporations of the pBluescript phleomycin A/B insert into the *T.b.brucei* neo A/B transformants have produced exceedingly slow growing double drug resistant transformants (Pools B, C, H, I, L, M, N) which have struggled for four weeks. The characterization of the drug resistant pools, is beyond the time scope of this Masters thesis and will be continued in Dr. T.W. Pearson's lab.

Characterization of the neomycin/phleomycin double drug resistant trypanosome pools. Before dilution cloning of the doubly drug resistant pools, the mutant phenotypes and nature of the insertion events within the heterologous population will be analyzed. KMP-11 immunofluorescence patterns of each stable transformant pool will be determined using the anti-KMP-11 mAb L157 to determine if individual trypanosomes within the population that are devoid of KMP-11. Southern blot analysis and genomic PCR will allow genetic characterization of the pools. Specific PCR reactions using a phleomycin forward primer and external diagnostic primer will indicate whether the insert is localized to the KMP-11 locus. Southern blot analysis similar to that performed on the neomycin drug resistant transformants will give further information about the genotypes found in the phleomycin resistant pools.

Isolation and characterization of individual double drug resistant transformants. Once dilution cloning of the promising pools of transformants has been performed, the individual clones can be characterized. The characterization of the these clones will encompass the methods previously describe for transformant pools, but will additionally include electron microscopy of the KMP-11 null mutants to illustrate morphological alterations with respect to the wild type trypanosomes.

DISCUSSION

Once the KMP-11 locus had been fully characterized, efforts were completely directed towards gene deletion experiments in the hope of elucidating the function of KMP-11. It was clear that the deletion would have to be of the four closely linked genes found in the KMP-11 locus. Because there were two alleles of the KMP-11 locus, two different deletion cassettes were used to sequentially delete the two loci. Integration of foreign DNA into the trypanosome genome occurs almost exclusively by homologous recombination and evidence from the literature suggested that a relatively small amount of homologous DNA was required for targeting of this event. The 5' and 3' flanking sequences of the KMP-11 locus provided a total of between 1200 bp (A/B constructs) and 1300 bp (A'/B constructs) of homologous DNA sequence, quite sufficient to target the deletion cassettes to the KMP-11 locus. Each set of cassettes was designed to tackle specific integration possibilities. The A'/B construct included 102 nucleotides of 5' untranslated upstream DNA that is repeated in front of each of the four genes of the KMP-11 locus. These stretches of conserved DNA contained poly-pyrimidine rich elements that are thought to be associated with *in vivo* splicing [Schurch *et al.*, 1994]. To safe-guard against a situation in which no functional mRNA transcript was produced due to faulty splicing of the pre-mRNA, the A'/B construct included this apparently essential sequence. Since the repeated region generates relatively large regions of homology upstream of each KMP-11 gene, the possibility existed that the A'/B constructs would integrate internally and only partially delete the locus. Thus, another set of constructs, A/B, that lacked this upstream conserved region was designed to address this problem.

The first round of knockouts using the neomycin drug resistance deletion cassettes generated two distinct sets of transformants. The *T.b.brucei* neo A/B transformants showed an integration pattern consistent with the complete deletion of one of the two KMP-11 gene loci, while the *T.b.brucei* neo A'/B transformants exhibited only a partial knockout of the KMP-11 locus, leaving a single copy of KMP-11 in close proximity to the deletion cassette (Figure 28, 29 & 30). The second round of gene knockouts was attempted using

an almost identical set of deletion cassettes in which the phleomycin drug resistance marker had replaced its neomycin counterpart. Several pools of transformants from the second set of electroporation have recently regrown to healthy densities. Immunofluorescence identification of the KMP-11 molecule in these populations indicated that the majority of the trypanosomes found in the pools still express KMP-11. Several of the transformant pools that did not regrow (pools B, C, H, I and L) contained transformants that survived for over 4 weeks in the presence of phleomycin, long after the time when the negative control population (electroporated without DNA) had died off (Table 3). The extended viability of these cultures in absence of proliferation may indicate that while full knockouts of KMP-11 genes are possible, the effect of such deletions renders the organism incapable of division and of prolonged survival. Whatever the reason for my inability to obtain proliferating KMP-11 deletion mutants, it would seem that KMP-11 is an essential molecule for the survival of African trypanosomes because the wide variety of conditions used in the knockout studies did not produce null mutants. The growth curve of the trypanosomes was analyzed and several time points throughout the curves were used for electroporation. Two distinct sets of constructs previously shown to integrate at the KMP-11 locus were used for the second round of knockouts. More than twenty-five separate experiments were performed with varying electroporation conditions (the voltage, the number of pulses and the amounts and type of insert DNA used). Genetic analysis of the transformants from the second round of knockouts will indicate the type of integration events that have occurred and give insight about the number of KMP-11 gene copies needed to produce sufficient levels of KMP-11 for survival of these African trypanosomes.

GENERAL DISCUSSION

The kinetoplastid membrane protein-11 was first identified in *Leishmania donovani* and was subsequently found to have a ubiquitous distribution among other kinetoplastids, including African trypanosomes. This amphipathic membrane associated molecule shows stage specific patterns of expression and is estimated to be one of the most abundant membrane molecules found in *T.brucei*. Initial Southern blot evidence indicated that KMP-11 was present as a single copy gene in a *T.b.rhodesiense* clone [Stebeck *et al.*, 1996]. Detailed analysis of *T.b.brucei* clone 427.01 showed that the KMP-11 protein is encoded by a tandem array of four tightly packed repeats of the KMP-11 gene. The elucidation of the highly repetitive KMP-11 locus relied on the production and sequencing of a high redundancy randomly disrupted sequencing library. The alignment of overlapping sequences from this library allowed the identification of a small non-identical region within the sequence, which was a key to assembling the entire locus. The tandem array of the KMP-11 gene cluster is a motif commonly found in trypanosome genetic organization. A unique and interesting aspect of the KMP-11 locus was that the complete sequence provided an insight into the duplication events that likely occurred during its evolutionary.

With the genetic foundations laid, KMP-11 gene deletion studies ensued. The first round of KMP-11 knockouts used the neomycin drug resistance cassettes and deleted one of the two KMP-11 allelic loci, producing +/null mutants. The lack of KMP-11 null/null mutants from the second round of knockouts indicates that KMP-11 is required for trypanosome survival. Supporting evidence has come from work done by Armando Jardim (University of Oregon) in which KMP-11 gene deletions in *Leishmania donovani* failed to produce viable null mutants (personal communication). Whether the necessity for KMP-11 stems from a role in the division process of the parasite or from other critical roles is unresolved. For example, KMP-11's disposition in the flagellar pocket allows the speculation that its removal may interfere with the endocytotic processes that occurs in this location. Thus, deletion of this molecule could halt or impair transport of essential molecules, such as transferrin or glucose, crippling the parasite and causing the slow but

inevitable death seen in several of the transformant pools from the second electroporation. Further, KMP-11 could have an essential structural or functional role in the trypanosome membrane. KMP-11 shows tight association and co-isolation with the major membrane associated molecules, lipophosphoglycan (LPG) and procyclin, in *L.donovani* and African trypanosomes respectively. Both LPG in *Leishmania* and procyclin in trypanosomes have been purported to form a carbohydrate rich calix, which surrounds and protects the parasites. *L. donovani* KMP-11 has been suggested to stabilize LPG within the parasite membrane by regulating the overall lipid bilayer pressure via a membrane association with LPG [Jardim *et al.*, 1995]. The relationship between KMP-11 and the major membrane molecule of PCF African trypanosomes, procyclin, mirrors that seen in *Leishmania.* Procyclin and KMP-11 are co-isolated by reverse-phase HPLC and share similar patterns of expression, suggesting that KMP-11 may have a similar stabilizing role for procyclin in the membrane of African trypanosomes. While the true role of KMP-11 remains unresolved, the lack of null mutants from the deletion studies suggests that this ubiquitous and highly conserved membrane molecule is essential to the survival of the African trypanosome.

Literature cited

Alexandre S., Guyaux M., Murphy N., Coquelet H., Pays A., Steinert M., Pays E. 1988. Putative genes of a variant-specific antigen gene transcription unit in *Trypanosoma brucei*. Molecular and Cellular Biology. 8: 2367-2378.

Amar B., Pays E., Tebabi P., Dero B., and Seebeck T. 1988. Structure and transcription of the actin gene of *Trypanosoma brucei*. Molecular and Cellular Biology. 8: 2166-2176.

Barrett M., Tetaud E., Seyfang A., Bringaud F., Baltz T. 1995. Functional expression and characterization of the *Trypanosoma brucei* procyclic glucose transporter, THT2. Biochemical Journal. 312: 687-691.

Bastin P., Coppens I., Saint J., Baudhuin P., Opperdoes F., and Courtoy P. 1994. Identification of a specific epitope on the extracellular domain of the LDL-receptor of *Trypanosoma brucei brucei*. Molecular and Biochemical Parasitology. 63: 193-202.

Bastin P., Stephan A., Jayne R., Saint J., Opperdoes F., and Courtoy P. 1996. An Mr 145, 000 low-density lipoprotein (LDL)-binding protein is conserved throughout the *Kinetoplastida* order. Molecular and Biochemical Parasitology. 76: 43-56.

Benne R., Van den Burg J., Brakenhoff J., Sloof P., Van Boom J., Tromp M. 1986. Major transcript of the frame shifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell. 46: 819-826.

Benne R. 1994. RNA editing in trypanosomes. European Journal of Biochemistry. 221: 9-23.

Berberof M., Pays A., and Pays E. 1991. A similar gene is shared by both the variant surface glycoprotein and procyclin gene transcription units of *Trypanosoma brucei*. Molecular and Cellular Biology. 11: 1473-1479.

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

Bodenteich A., Chissoe S., Wang F., Roe B., Rowen L. and Koop B. Adams MD, Fields C, Vener JC, (editors). In: Automated DNA sequencing and analysis. London: Academic Press, 1994; 42-49 and 167-174.

Borst P. 1986. Discontinuous transcription and antigenic variation in trypanosomes. Annual Reviews in Biochemistry. 55: 701-732.

Brickman M. and Balber A. 1993. *Trypanosoma brucei rhodesiense*: membrane glycoproteins localized primarily in endosomes and lysosomes of bloodstream forms. Experimental Parasitology. 76: 329-344.

Bridge M., Zhou Q., Koop.B., and Pearson T. 1998. Cloning and characterization of the kinetoplastid membrane protein-11 locus of *Trypanosoma brucei*. Molecular and Biochemical Parasitology. 91: 359-363.

Birnboim H. and Doly J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Research. 7: 1513-1523.

Bringaud F. and Baltz T. 1992. A potential hexose transporter gene expressed predominantly in the bloodstream form of *Trypanosoma brucei*. Molecular and Biochemical Parasitology. 52: 111-121.

Bringaud F. and Baltz T. 1993. Differential regulation of two distinct families of glucose transporter genes in *Trypanosoma brucei*. Molecular and Cellular Biology. 13: 1146-1154.

Brundell P., Rudenko G., and Borst P. Targeting of exogenous DNA into *Trypanosoma brucei* requires a high degree of homology between donor and target DNA. Molecular and Biochemical Parasitology. 76: 215-229.

Butikofer P., Ruepp S., Boschung M. and Roditi I. 1997. GPEET procyclin is the major surface glycoprotein of *Trypanosoma brucei brucei* procyclic culture forms. Biochemical Journal. 326: 415-423.

Carrington M., and Boothroyd J. 1996. Implications of a conserved structural motif in disparate trypanosome surface proteins. Molecular and Biochemical Parasitology. 81: 119-126.

Carruthers V., Navarro M. and Cross G. 1996. Targeted disruption of expression site-associated gene-1 in bloodstream-form *Trypanosoma brucei*. Molecular and Biochemical Parasitology. 81: 65-79.

Cattaneo R. 1992. RNA editing: in chloroplasts and brain. Trends in Biochemical Science. 17: 4-5.

Clayton C. 1992. Developmental regulation of nuclear gene expression in *Trypanosoma brucei*. Progress in Nucleic Acid Research and Molecular Biology. 43: 37-66.

Coppens I., Bastin P., Courtoy P., Baudhuin P., Opperdoes F. 1991. A rapid method purifies a glycoprotein of Mr 145,000 as the LDL receptor of *Trypanosoma brucei brucei*. Biochemical and Biophysical Research Communications. 178: 185-191.

Coppens I., Levade T. and Courtoy P. 1995. Host plasma low density lipoprotein particles as an essential source of lipids for the bloodstream forms of *Trypanosoma brucei*. Journal of Biological Chemistry. 270: 5736-5741.

Davies J., and Smith D. 1978. Plasmid-determined resistance to antimicrobial agents. Annual Reviews in Microbiology. 23: 469-518.

Deng C., and Capecchi M. 1992. Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and target locus. Molecular and Cellular Biology. 12: 3365-3371.

Donelson J., Hill K., and El-Sayed M. 1998. Multiple mechanisms of immune evasion by African trypanosomes. Molecular and Biochemical Parasitology. 91: 51-66.

Dower W., Miller J., Ragsdale C. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Research. 16: 6127-6145.

Drocourt D., Calmels T., Reynes J., Baron M., and Tiraby G. 1990. Cassettes of the *Streptoalloteichus hindustanus ble* gene for transformation of lower and higher eukaryotes to phleomycin resistance. Nucleic Acids Research. 18: 4009.

Eid J., and Webb B. 1991a. Homologous recombination in the tandem calmodulin genes of *Trypanosoma brucei* yields multiple products: compensation for deleterious deletions by gene amplification. Genes and Development. 2024-2032.

Eid J., and Webb B. 1991b. Stable integrative transformation of *Trypanosoma brucei* that occurs exclusively by homologous recombination. Proceedings of the National Academy of Science U.S.A.. 88: 2118-2121.

Ferguson M., Murray P., Rutherford H., and McConville M. 1993. A simple purification of procyclic acid repetitive protein and demonstration of a sialated glycosyl-phosphatidylinositol membrane anchor. Biochemistry Journal. 291: 51-55.

Gatignol A., Durand H., and Tiraby G. 1988. Bleomycin resistance conferred by a drug-binding protein. FEBS letters. 230: 171-175.

Greenwood, B. 1982. In Cecil's Textbook of Medicine (J.B. Wyngarden and L.H. Smith, eds.), Saunders, Philidelphia. p. 1725.

Haag J., O'hUigin C., and Overath P. 1998. The molecular phylogeny of trypanosomes: evidence for an early divergence of Salivaria. Molecular and Biochemical Parasitology. 91: 37-49.

Hager K. And Hajduk S. 1994. Mechanism of resistance of African trypanosomes to cytotoxic human HDL. Nature. 385: 823-826.

Hajduk S., Hager K. And Esko J. 1994. Human high density lipoprotein killing of African Trypanosomes. Annual Review of Microbiology. 48: 139-162.

Hanahan D. 1983. Studies on transformation of Escherichia coli with plasmids. Journal of Molecular Biology. 166: 557-580.

Hehl A., Vassella E., Braun R., Roditi I. 1994. A conserved stem-loop structure in the 3' untranslated region of procyclin mRNAs regulates expression in *Trypanosoma brucei*. Proceedings of the National Academy of Science U S A.. 1: 370-374.

Hug M., Hotz HR., Hartmann C., and Clayton C. 1994. Hierarchies of RNA-processing signals in a trypanosome surface antigen mRNA precursor. Molecular and Cellular Biology. 14 : 7428-7435.

Imboden M., Laird P., Affolter M., and Seekbeck T. 1987. Transcription of the intergenic regions of the tubulin gene cluster of *Trypanosoma brucei*: evidence for a polycistronic transcription unit in a eukaryote. Nucleic Acids Research. 15: 7357-7368.

Jackson D., Windle H., and Voorheis P. 1993a. The indentification, purification, and characterization of two invariant surface glycoproteins located beneath the surface coat barrier of bloodstream forms of *Trypanosoma brucei*. Journal of Biological Chemistry. 268: 8085-8095.

Jackson D., Smith D., Luo C. and Elliot J. 1993b. Cloning of a novel surface antigen from the insect stages of *Trypanosoma brucei* by expression in Cos cells. Journal of Biological Chemistry. 268: 1894-1900.

Jardim A. Tolson D., Turco S., Pearson T., and Olafson R. 1991. The *Leishmania donovani* lipophosphoglycan T-lymphocyte reactive component is a tightly associated protein complex. Journal of Immunology. 147:3538-43.

Jardim A. Funk V., Caprioli R. and Olafson R. 1995. Isolation and structural characterization of the *Leishmania donovani* kinetoplastid membrane protein-11 - a major immunoreactive membrane glycoprotein. Journal of Biochemistry. 305:307-313.

Jasin M., Elledge S., Davis R., Berg P. 1990. Gene targeting at the human CD4 locus by epitope addition. Genes and Development. 4: 157-166.

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

Kelly J., Ward H., Miles M., and Kendall G. 1992. A shuttle vector which facilitates the expression of transfected genes in *Trypanosoma cruzi* and Leishmania. Nucleic Acids Research 20: 3960 - 3969.

Koenig E., Delius H., Carrington M., Williams R., Roditi I. 1989. Duplication and transcription of procyclin genes in *Trypanosoma brucei*. Nucleic Acids Research. 17: 8727-8739.

Koenig E., Yamage M., and Roditi I. 1992. A procyclin-associated gene in *Trypanosoma brucei* encodes a polypeptide related to ESAG 6 and 7 proteins. Molecular and Biochemical Parasitology. 55: 135-146.

Landweber L., and Gilbert W. 1993. RNA editing as a source of genetic variation. Nature 363: 179-182.

Langreth S. and Balber A. 1975. Protein uptake and digestion in bloodstream and culture forms of *Trypanosoma brucei*. Journal of Protozoology. 22: 40-53.

Layden R., and Eisen H. 1988. Alternative trans-splicing in *Trypanosoma equiperdum*: Implications for splice site seletion. Molecular and Cellular Biology. 8 :1352-1360.

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

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

Lee M., and van der Ploeg L. 1997. Transcription of protein-coding genes in trypanosomes by RNA polymerase I. Annual Review of Microbiology. 51: 463-489.

Ligtenber M., Bitter W., Kieft R., Steverding D., Janssen H., Calafat J and Borst P. 1994. Reconstruction of a surface transferrin binding complex in insect form *Trypanosoma brucei*. The EMBO Journal. 13; 11:2565-2573.

Lips S., Revelard P. and Pays E. 1993. Identification of a new expression site-associated gene in the complete 30.5 kb sequence from the AnTat 1.3A variant surface protein gene expression site of *Trypanosoma brucei*. Molecular and Biochemical Parasitology. 62: 135-138.

Lucas R., Magez S., De Leys R., Frensen L., Scheerlinck J., Rampelberg M., Sablon E. and De Baetselier P. 1994. Mapping the lectin-like activity of tumor necrosis factor. Science. 263: 814-817.

Manivasakam P., Weber S., McElver J., and Schiestl R. 1995. Micro-homology mediated PCR targeting in *Saccharomyces cerevisiae*. Nucleic Acids Research. 23: 2799-2800.

Maudlin I., and Welburn S. 1994. Maturation of Trypanosome infection in tsetse. Experimental Parasitology. 79: 202-205.

McLaughlin J. 1987. **Trypanosoma rhodesiense**: antigenicity and immunogenicity of flagellar pocket membrane components. Experimental Parasitology. 64: 1-11.

Molyneux H., and Ashford W. 1983. The biology of *Trypanosoma* and *Leishmania*, parasites of man and domestic, Taylor and Francis, London.

Mowatt M., and Clayton C. 1987. Development regulation of a novel repetitive protein of *Trypanosoma brucei*. Molecular and Cellular Biology. 8: 2833-2844.

Nolan D., Jackson D., Windle H., Pays A., Geuskens M., Michel A., Voorheis H., Pays E. 1997. Characterization of a novel, stage-specific, invariant surface protein in *Trypanosoma brucei* containing an internal, serine-rich, repetitive motif. Journal of Biological Chemistry. 272: 29212-29221.

Navarro M. and Cross G. 1996. DNA rearrangements associated with multiple consecutive directed antigenic switches in *Trypanosoma brucei*. Molecular and Cellular Biology. 16: 3615-3625.

Olsson T., Bakhiet M., Hojeberg B., Ljungdahl A., Edlund C., Andersson G., Ekre H., Funf leung W., Mak T., Wigzell H., Fiszer U. And Kristensson K. 1994. CD8 is critically involved in lymphocyte activation by a *T. brucei brucei* - released molecule. Cell. 72: 715-727.

Ojala D., Crews S., Montoya J., Gelfand R., Attardi G. 1981. A small polyadenylated RNA (7 S RNA), containing a putative ribosome attachment site, maps near the origin of human mitochondrial DNA replication. Journal of Molecular Biology. 150: 303-314

Orr-Weaver., Szostak J., and Rothstein R. 1981. Yeast transformation: A model system for the study of recombination. Proceedings of the National Academy of Science U.S.A.. 78: 6345-6358.

Pays E. Tebabi P., Pays A., Coquelet H., Revelard P., Salmon D., and Steinert M. 1989. The genes and transcripts of an antigen gene expression site from *T. brucei*. Cell. 57: 835-845.

Pays E. Vanhamme L., and Berberof M. 1994. Genetic controls for the expression of surface antigens in African trypanosomes. Annual Reviews in Microbiology. 48: 25-52.

Pays E. and Nolan D. 1998. Expression and function of surface proteins in *Trypanosoma brucei*. Molecular and Biochemical Parasitology. 91: 3-36.

Pierce J., Sauer B. and Sternberg N. 1992. A positive selection vector for cloning high molecular weight DNA by the bacteriphage P1 system: improved cloning efficacy. Proceedings of the National Academy of Sciences U.S.A. 89: 2056-2060.

Pearson T. In Lackmann P., Peters K., Rosen F., and Walport M. 1993. In: Clinical aspects of immunology: The immunology of African Trypanosomes. 81: 1599-1612.

Povinelli C. and Gibbs R. 1992. Large scale sequencing library production: an adaptor-based strategy. Analytical Biochemistry. 210:16-26.

Roditi I., Schwarz H., Pearson T., Beecroft R., Liu M., Richardson J., Buhring H., Pleiss J., Bulow R., Williams R. And Overath P. 1989. Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of *Trypanosoma brucei*. Journal of Cell Biology. 108: 737-746.

Roditi I. and Pearson T.W. 1990. The procyclin coat of African trypanosomes (or the not-so-naked trypanosome). Parasitology Today. 6: 79-81

Roditi I., Furger A., Ruepp S., Schurch N., and Butikofer P. 1998. Unravelling the procyclin coat of *Trypanosoma brucei*. Molecular and Biochemical Parasitology. 91: 117-130.

Rolin S., Hanocq-Quertier J., Paturiaux-Hanocq F., Nolan D., Salmon D., Webb H., Carrington M., Voorheis P., Pays E. 1996. Simultaneous but independent activation of adenylate cyclase and glycosylphosphatidylinositol-phospholipase C under stress conditions in *Trypanosoma brucei*. Journal of Biological Chemistry. 271: 10844-10852.

Sambrook J., Fritsch F., and Maniatis T. 1989. In Molecular Cloning: A Laboratory Manual. Second Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor , NY.

Schell D., Evers R., Preis D., Ziegelbauer K., Kiefer H., Lottspeich F., Cornelissen A., Overath P. 1991. A transferrin-binding protein of *Trypanosoma brucei* is encoded by one of the genes in the variant surface glycoprotein gene expression site. EMBO Journal. 10: 1061-1066.

Schurch N., Hehl A., Vassella E., Brun R., and Roditi I. 1994. Accurate polyadenylation of procyclin mRNAs in *Trypanosoma brucei* is determined by pyrimidine-rich elements in the intergenic regions. Molecular and Cellular Biology. 14: 3668-3675.

Shapiro SZ, Naessens J, Liesegang B, Moloo SK, Magondu J. 1984. Analysis by flow cytometry of DNA synthesis during the life cycle of African trypanosomes. Acta Tropica. 41: 313-323.

Shapiro S. and Pearson T.W. In: Pearson T. (editor). 1986. African Trypanosomiasis: Antigens and Host-Parasite Interactions. 215-274.

Stebeck C. and Pearson T.W. 1994. Major surface glycoproteins of procyclic stage African Trypanosomes. Experimental Parasitology. 78: 432-436.

Stebeck C., Beecroft R., Singh B., Jardim A., Olafson R., Tuckey C., Prenevost K. and Pearson T.W. 1995. Kinetoplastid membrane protein-11 (KMP-11) is differentially expressed during the life cycle of African trypanosomes and is found in a wide variety of kinetoplasid protozoan parasites. Molecular and Biochemical Parasitology. 71:1-13.

Stebeck C., Frevert U., Vassella E., Roditi I., and Pearson T.W. 1996. Molecular characterization of glycosomal NAD^+-dependent glycerol-3-phosphate dehydrogenase from *Trypanosoma brucei rhodesience.* Molecular and Biochemical Parasitology. 79: 145-148.

Stebeck C., Baron G., Beecroft R., and Pearson T. 1996. Molecular characterization of the kinetoplastid membrane protein-11 from African trypanosomes. Molecular and Biochemical Parasitology. 81:81-88.

Steverding D., Stierhof Y., Chaudhri M., Lightenberg M., Schell D., Beck-Sickinger A., and Overath P. 1994. ESAG 6 and 7 products of *Trypanosoma brucei* form a transferrin binding protein complex. European Journal of Cell Biology. 64: 78-87.

Sonnhammer E. and Durbin R. 1995. A dot matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene. 167: 1-10.

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

ten Asbroek A., Mol C., Kieft R. and Borst P. 1993. Stable transformation of *Trypanosoma brucei.* Molecular and Biochemical Parasitology. 59: 133-142.

Turner C., Sternberg J., Buchanan N., Smith E., Hide G., Tait A. 1990. Evidence that the mechanism of gene exchange in *Trypanosoma brucei* involves meiosis and syngamy. Parasitology. 101: 377-386.

Tolson D., Turco S., Beecroft R., Pearson T. 1989. The immunochemical structure and surface arrangement of *Leishmania donovani* lipophosphoglycan determined using monoclonal antibodies. Molecular and Biochemical Parasitology. 35: 109-118.

Van der Ploeg L., Valerio D., De Lange T., Bernards A., Borst P., Grosveld F. 1982. An analysis of cosmid clones of nuclear DNA from *Trypanosoma brucei* shows that the genes for variant surface glycoproteins are clustered in the genome. Nucleic Acids Research. 10: 5905-5923.

Vanhamme L., and Pays E. 1995. Control of gene expression in trypanosomes. Microbiological Reviews. 233-240.

Vassella E., Braun R. and Roditi I. 1994. Control of polyadenylation and alternative splicing of transcripts from adjacent genes in a procyclin expression site: a dual role for polypyrimidine tracts in trypanosomes? Nucleic Acids Research. 22:1359-1364.

Vickerman K. 1962. Transcripts of the Royal Society of Tropical Medicine Hyg. 56:487.

Vickerman K. 1965. Polymorphism and mitochondiral activity in sleeping sickness trypanosomes. Nature. 208: 762-766.

Vickerman K. 1976. In: Biology of the Kinetoplastida. W.H.T. Lumsden and D.A. Evans, eds. Academic Press, London. p. 1-76.

Vickerman K. 1989. Trypanosome sociology and antigenic variation. Parasitology Supplement S37- S47.

Vijayasarathy S, Ernest I, Itzhaki JE, Sherman D, Mowatt MR, Michels PA, Clayton CE. 1990. The genes encoding fructose bisphosphate aldolase in *Trypanosoma brucei* are interspersed with unrelated genes. Nucleic Acids Research. 18: 2967-2975.

Webster P. and Grab D. 1988. Intracellular colocalization of variant surface glycoprotein and transferrin-gold in *Trypanosoma brucei*. Journal of Cell Biology. 106: 279-288.

Webb J., and McMaster R. 1994. Leishmania major HEXBP deletion mutants generated by double targeted gene replacement. Molecular and Biochemical Parasitology. 63: 231-242.

WHO Publications. 1979. The African Trypanosomiases: Report of a joint WHO expert committee and FAO expert consultation. WHO: Geneva 1-96.

Ziegelbauer K., Multhaup G. and Overath P. 1992. Molecular characterization of two invariant surface glycoproteins specific for the bloodstream stage of *Trypanosoma brucei*. The Journal of Biological Chemistry. 267: 10797-10803.

Ziegelbauer K. and Overath P. 1992. Identification of invariant surface glycoproteins in the bloodstream stage of *Trypanosoma brucei*. Journal of Biological Chemistry. 267: 10791-10796.

Ziegelbauer K., Rudenko G., Kieft R. and Overath P. 1995. Genomic organization of an invariant surface glycoprotein gene family of *Trypanosoma brucei*. Molecular and Biochemical Parasitology. 69: 53-63.

Zimmer C. 1998. A sleeping storm. Discovery Magazine. 87-94.

Zomerdijk J., Kieft R., Duyndam M., Shiels P., Borst P. 1991. Antigenic variation in *Trypanosoma brucei*: a telomeric expression site for variant-specific surface glycoprotein genes with novel features. Nucleic Acids Research. 19: 1359-1368.

VITA

Surname: Bridge Given Names: Michael Adrian Anthony

Place of Birth: Cheltenham, England Birth Date: January 02, 1973

Educational Institutions Attended:

University of Victoria 1991-1995

Degrees Awarded:

B.Sc. (Major) University of Victoria 1995

Honors and Awards:

Dr. Julius Schleicher Graduate Scholarship	1997-1998
The Ray Hadfield Memorial Fellowship	1996-1997
President's Research Scholarship	1996-1998
Natural Sciences and Engineering Research Council of Canada (NSERC)	1996-1998

Publications:

Michael A. Bridge, Qun Zhou, Ben F. Koop and Terry W. Pearson. 1998. Cloning and Characterization of the Kinetoplastid Membrane Protein-11 Gene Locus of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*. 91: 359-363.

Abstracts and Presentations Published:

Michael A. Bridge, Qun Zhou, Ben F. Koop and Terry W. Pearson. 1997. Cloning and Characterization of the Kinetoplastid Membrane Protein-11 Gene Locus of *Trypanosoma brucei*. The Seattle Protozoology Conference. March, 1997. Seattle Pacific University. Seattle, Washington, USA.

PARTIAL COPYRIGHT LICENSE

Title of Dissertation: **The Identification and Characterization of the kinetoplastid Membrane Protein-11 (KMP-11) Gene Locus from Trypanosoma Brucei and its Attempted Genetic Deletion**

Author: [redacted]

Michael A. A. Bridge

Date: Dec 18th / 1998