The Role of Ku In Antigenic Variation, DNA Repair And Telomere Maintenance In African Trypanosomes

Colin Conway
Wellcome Centre for Molecular Parasitology
The Anderson College
University Of Glasgow

Submitted for the degree of Doctor of Philosophy

April, 2002

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Declaration

I declare that this thesis and the results presented in it are entirely my own work except where otherwise stated.

Colin Conway

ACKNOWLEDGMENTS

I would like to thank my supervisor, Dave Barry, for the continual support and encouragement through the progression of this project. I would additionally like to thank all my colleagues especially the "baboons", for their assistance and advice over the years and in particular I would like to acknowledge Richard McCulloch for constant assistance. Thanks a lot to Pete and Mary for trawling through this document prior to submission. Credit must also be given to Ces, Tom and Daidín for putting up with me over the years and Neil Lennon for being such a good pal.

I would finally like to thank members of Keith Gull's laboratory, in particular Klaus Ersfeld, for assistance with FISH analysis and PFGE conditions.

This work was supported by a studentship from Wellcome Trust.

This thesis is dedicated to the memory of my mother.

CONTENTS

FIGURES LIST

CHAPTER 1

CHAPTER 3

CHAPTER 4

TABLES LIST

ABBREVIATIONS

BAC	bacteria artificial chromosome
BC	basic copy
BES	bloodstream expression site
BIR	break-induced replication
bp	base pairs
DNA-PK_{cs}	DNA dependent protein kinase catalytic subunit
DSB	double strand breaks
ELC	expression linked copy
ES	expression site
*ESAG*s	expression site associated genes
EST	expressed sequence tags
FISH	fluorescent in situ hybridisation
gDNA	genomic DNA
GPI	glycophosphatidylinositol
GSS	genome survey services
GUTat	Glasgow University Trypanozoon antigen type
HTG	high throughput genome
Ig	immunoglobulin
ILTat	International Laboratory for research on animal diseases, Trypanozoon antigen type
kb	kilobases
MES	metacyclic *VSG* expression site
MITat	Minnesota Trypanozoon antigen type
MMS	methyl methanesulphonate
MVSG	metacyclic variant surface Glycoprotein
NBS	nijmegan breakage syndrome
NCBI	National Centre for Biotechnology Information
NHEJ	non-homologous end joining
NPC	nucleoporin complex

ORF	open reading frame
PFGE	pulsed field gel electrophoresis
PHR	primary homology region
rRNA	ribosomal RNA
RT	room temperature
TIGR	the Institute for Genomic Research
TLM	telomere length maintenance
TPE	telomere position effect
TREU	Trypanosomiasis Research Edinburgh University
UTR	untranslated region
VSG	variant surface glycoprotein

ABSTRACT

The process of antigenic variation in African trypanosomes allows the survival of the parasite by constantly switching the variant surface glycoprotein (VSG) expressed in their surface. There are believed to be several hundred copies of these silent *VSG* genes in the parasite's genome and they are expressed differentially. The majority of these genes are not capable of being transcribed *in situ* and must therefore be expressed from specialised transcriptional units known as bloodstream expression sites (*BES*s). Only one such site is active at any one time, ensuring that a single VSG is expressed in the trypanosome's surface coat. Switching the expressed VSG involves replacing the *VSG* in the active *BES*, or activating a new *BES* in conjunction with silencing the previously active.

Differential expression of variant surface glycoprotein (VSG) genes, has a strong association with telomeres. All *BES*s are telomeric and differential activation involves recombination into the telomeric environment or silencing/activation of subtelomeric promoters. A number of pathogen contingency gene systems associated with immune evasion involve telomeric loci, which has prompted speculation that chromosome ends provide conditions conducive for the operation of rapid gene switching mechanisms. Ku is a protein associated with yeast telomeres that is directly involved in DNA recombination and gene silencing. The main aim of this thesis was to test the hypothesis that Ku in trypanosomes is centrally involved in differential *VSG* expression. In order to compare trypanosome Ku homologues with those from other organisms, it was necessary to compile homology alignments with other Ku homologues using Clustal W analysis. This task was made easier in recent years with definition of the most homologous portions of each protein as primary homology regions (PHRs) (Gell & Jackson, 1999).

Mutants were developed lacking both alleles of each putative homologue partner of the *KU* heterodimer (*KU70* and *KU80*). Initial phenotypic analysis meant adapting assays used in organisms in which Ku had previously been characterised to determine whether the genes cloned using trypanosome genome databases were indeed homologues. We looked at telomere length maintenance (TLM), DNA double-strand break repair efficacy and telomere localisation in *ku* mutants developed in a bloodstream form trypanosome

strain manipulated to test *VSG* switching frequencies. Following determination of homologue status through phenotypic analysis and sequence alignment comparisons, assays were used to determine whether the mutants were handicapped in their ability to switch the surface coat being expressed in the bloodstream stage of the parasite. We subsequently tested the ability of procyclic form mutants to maintain previously inactive metacylic expression sites (*MES*s) silent. We showed that, despite trypanosome Ku having a close involvement in TLM, there is no difference between *KU* wild-type and null mutants in the rate of *VSG* switching or in the silencing of the telomeric promoters of the *VSG* subset that is expressed in the tsetse fly. We also showed no effect on the frequency of *VSG* genes being switched in the bloodstream form stage of the parasite. This argues against the theory that telomeric location is directly associated with differential contingency gene expression.

Our analysis meant attempting to develop an assay to analyse DNA repair through non-homologous end joining (NHEJ) of *ku* null mutants. Ku is central in catalysis of NHEJ in both yeasts and mammalian cells. This is a relatively simple form of repair with increased importance in higher eukaryotic DNA repair systems. NHEJ, however, has not previously been demonstrated in African trypanosomes. Initial attempts to define a role for NHEJ in trypanosome DNA repair involved crude plasmid rejoining assays whereby restriction enzyme digested plasmids were used to transform trypanosome populations. DNA was subsequently prepared from these populations and the preparations were used to transform competent bacterial cells. The number of bacterial transformants would then represent the ability of trypanosome populations to repair naked DNA substrates by NHEJ. Further attempts to monitor the role of NHEJ in trypanosomes meant the development of more complex assays. Subsequent experiments looked at the fate of exogenously introduced restriction enzyme target sites after transient transformation with cassettes encoding the restriction enzyme.

A final analysis looked for the presence of NHEJ in homologous recombination-deficient trypanosomes. Disrupting this element of DNA repair would hopefully lead to other forms of repair becoming detectable, and even up-regulated. Rad51, in yeast a member of the Rad52 epistasis group (integral in yeast homologous recombination), had

previously been demonstrated to be involved in DNA repair in trypanosomes (McCulloch & Barry, 1999). *rad51* mutants were electroporated with cassettes containing non-compatible ends that would prevent their integration into the endogenous genome via conventional homologous recombination. This cassette also contained promoter DNA sequence to allow selection in the event of integration into non-transcribed regions of the genome. Study of the junctions encompassing the integration sites of the cassette allowed investigation into how the cassettes were integrated, and revealed to us the extent of the sequence homology required to catalyse integration. The method of repair detection observed indicated that classical homologous recombination is not the only pathway utilised by African trypanosomes to metabolise DNA double-strand breaks.

CHAPTER 1

INTRODUCTION

1.1 General Introduction

Trypanosomes are unicellular eukaryotes of the order kinetoplastida. Several trypanosome species, including *Trypanosoma brucei*, are extracellular parasites, proliferating in the tissue fluids, capillary beds and vascular systems of mammalian hosts (Vickerman, 1985). Host infection results in Nagana in cattle and sleeping sickness in humans. *T. brucei* subspecies vary in host range and associated pathologies (Tait *et al.*, 1985; MacLeod, 1999). One such subgroup contains two host variants, *T. brucei brucei* and *T. brucei rhodesiense*. *T. b. brucei* causes Nagana in cattle and also infects a range of African wildlife but is, however, sensitive to human serum. *T. b. rhodesiense* is resistant to the lytic effect of human serum *in vitro* and is responsible for the acute form of human sleeping sickness. *T. b. gambiense*, the second subspecies, is also responsible for human sleeping sickness, causing the chronic form of the disease.

The life cycle of the African trypanosome is classically sub-divided into replicative and transmission stages (Fig 1). The replicative stages, including the long slender bloodstream form and the procyclic and epimastigote forms, undergo mitotic division. Each of these stages has the function of establishing infection through proliferation. All bloodstream form cells are covered by a characteristic surface coat composed of the VSG (variant surface glycoprotein) (Cross, 1975), which protects the parasite from host immune attack. Upon uptake by the tsetse fly, stumpy non-dividing bloodstream form cells develop into the procyclic form, which inhabits the midgut of the fly. Recent investigations have demonstrated that the number of parasites in the midgut drops substantially over the first few days following infection (Van Den Abbeele *et al.*, 1999). Concomitant with population development in the midgut, procyclic cells become morphologically elongated and move towards the anterior midgut (Van Den Abbeele *et al.*, 1999). Colonisation of the midgut culminates in the presence of a long trypomastigote form, referred to as mesocyclic cells (Vickerman, 1985). These occur in the midgut near the proventriculus (Vickerman, 1985; Van Den Abbeele *et al.*, 1999). This form then enters the proventriculus, changing into longer (post-mesocyclic) cells, which start to replicate DNA from a diploid to a tetraploid state, and migrate to the foregut and proboscis (Van Den Abbeele *et al.*, 1999). Post-mesocyclic cells differentiate

into long epimastigote cells (with no loss of genetic material) and it is suspected that this stage migrates to the salivary glands (Van Den Abbeele *et al.*, 1999). The long epimastigote cells divide asymmetrically into two morphologically distinct (diploid) daughter cells, the long and short epimastigote forms (Van Den Abbeele *et al.*, 1999). It is not clear what happens to the long epimastigote stage but it is thought to be a dead end in development. However, the short form attaches to the epithelial cells of the salivary glands where it multiplies and progresses to the nascent metacyclic stage (Vickerman, 1985). The parasites once again express a VSG coat, before transforming into infective metacyclic cells, which are ultimately transmitted back to the mammalian host when the fly feeds.

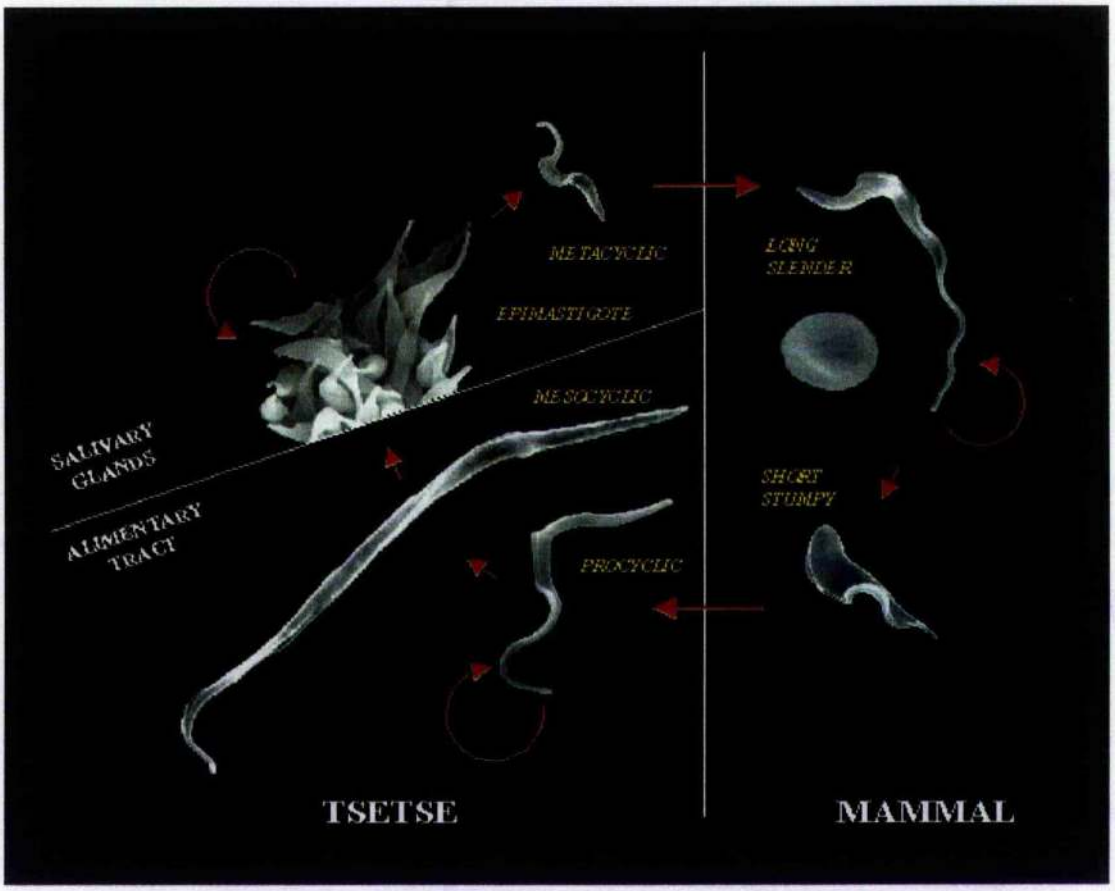

Figure 1. Life cycle of *Trypanosoma brucei*. Life cycle stages are represented as scanning electron micrograph images, reproduced to scale. Replicative stages of the life cycle are depicted using curved arrows. Epimastigote stage cells are shown as a cluster attached to the salivary glands of the tsetse fly. This diagram has been depicted previously in Barry & McCulloch, 2001. Scanning electron micrographs were provided courtesy of Dr Laurence Tetley, Electron Microscopy Centre, Institute of Biomedical and Life Sciences, University of Glasgow, Scotland, UK.

1.2 African trypanosome genome

There are three different size classes of nuclear chromosomes in the 35 Mb African trypanosome genome: the megabase chromosomes (1-6 Mb), the intermediate chromosomes (200-900 kb) and the minichromosomes (50-150 kb). The megabase

chromosomes are diploid (Hope *et al.*, 1999; Tait *et al.*,1989; Melville *et al.*, 1998), while the other two sub-classes are of uncertain ploidy. There are at least 11 pairs of megabase chromosomes. The sizes of chromosomal pairs and the homologues within the pair vary significantly within and among *T. brucei* stocks. These polymorphisms are believed to be due to rearrangements of *VSG* genes and *ES*s, retrotransposable elements and tandemly arrayed housekeeping genes and unknown DNA sequences. Sequence data released for chromosome I reveal large internal tracts of housekeeping and trypanosome specific genes. Telomeres are occupied by elements of both metacyclic and bloodstream expression sites (see below). Interestingly, expression sites appear to harbour haploid regions of the mainly diploid megachromosomes.

Minichromosomes are composed predominantly of repetitive sequence, containing tracts of 177-bp repeats. Indeed, these repeats constitute greater than 90% of the DNA sequence of some minichromosomes. GC- and AT-rich repeat sequences are found between the 177-bp repeats and telomeres (Weiden *et al.*, 1991). It is widely believed that minichromosomes serve as a repository for *VSG* genes. Activation of *VSG*s from minichromosomal loci requires transposition to an expression site on a megachromosome, and there are believed not to be any such expression cassettes present in minichromosomes. Two minichromosomes, of 55 and 60 kb, have been found to have a rRNA gene promoter near one telomere, perhaps because of a recombination with a repetitive rRNA gene locus in a megabase chromosome (Zomerdijk *et al.*, 1992). Insertion of a *neo*^{r} gene downstream of this promoter in the 55 kb molecule results in neomycin resistance. Transcription persisted through 130 generations without drug selection, indicating that the minichromosome is stably inherited during mitosis. Furthermore, Alsford *et al.*, (2001) have demonstrated, using DNA markers for specific minichromosomes, that minichromosomes are mitotically stable within the population over a period of more than 5 years. Intermediate chromosome size and number vary among stocks and they contain few, if any, unique markers or housekeeping genes (see El-Sayed *et al.*, 2000 for review). It is not known if this chromosome sub-division contains its own individual repeat sequences. The intermediate chromosomes do, however, contain telomere-linked *VSG*s or *VSG ES*s.

The African trypanosome genome network was initiated in 1995 in order to help elucidate the parasite's genome. Two organisations, The Institute for Genomic Research (TIGR) and the Sanger Centre, have undertaken a large scale sequencing strategy. The chosen clone in this project was *T. b. brucei* TREU 927 GUTat 10.1 expressing a specific VSG and derived from stock TREU (Trypanosomiasis Research Edinburgh University) 927/4 (GPAL/KE/70/EATRO1534). The TIGR strategy has two parts, which should yield 15 Mb of discontinuous single-pass sequence (50% of the non-minichromosomal genome) and ~6 Mb of completed sequence on selected chromosomes. The first phase of the strategy involves BAC, P1 and whole genome sheared plasmid DNA libraries. The second phase of the strategy involves highly accurate sequencing of specific *T. brucei* chromosomes by selecting minimally overlapping BACs for complete sequencing. All of these end sequences have been submitted to the GSS division of Genbank. Contiguous *T. brucei* BAC sequence data generated at TIGR are continuously deposited at various stages of completion to the High Throughput Genome (HTG) sequences database at the National Centre for Biotechnology Information (NCBI).

T. brucei chromosome I has been wholly sequenced using a chromosome shotgun approach. In April 2000, the Wellcome Trust, via its Beowulf Genomics initiative, awarded more funding to the *T. brucei* genome project at the Sanger Institute. The additional funding builds on the pilot project already under way (Chromosome Ia). The new project, a collaboration between the Sanger Institute, and the *T. brucei* Genome Network, will determine the complete sequences of chromosomes IX (3.5 Mb), X (4.4 Mb) and XI (5.2 Mb). In addition, the project includes end-sequencing of clones, thus contributing to the continuing process of gene discovery.

Completion of genome sequencing projects will allow analysis of the function and relationships of genes. The EST sequencing projects are extremely important in the development of any such understanding. This should ultimately provide insights into potential targets that may be used in the development of antiparasitic agents.

1.3 Antigenic Variation

T. brucei is capable of evading host immune responses by a process known as antigenic variation. Trypanosomes maintain host infection due to the ability of cells within the population to switch the surface coat they express. It is known that a single species of this coat protein, the VSG glycoprotein, is expressed at any one time (Cross, 1975). There are probably several hundred individual *VSG* genes in the trypanosome genome, with confirmation of this likely to result from the *T. brucei* genome project. Genes encoding these surface coat proteins are expressed from exclusive expression cassettes, termed expression sites, which are situated in sub-telomeric loci. *VSG*s are expressed during two separate stages of the life cycle, the metacyclic phase which occurs in the salivary glands of the tsetse fly vector and the bloodstream phase which occurs in the parasite's mammalian host. Several mechanisms are known to cause *VSG* switching, although protein complexes involved in catalysing reactions culminating in switching the expressed *VSG* have not yet been characterised.

Much work has been carried out recently to try to advance our knowledge of the molecular mechanisms underlying trypanosome antigenic variation. Parallels have been drawn between switching events occurring in both life cycle stages (bloodstream and metacyclic) with available evidence suggesting that there exists a common evolutionary origin for the two processes (Barry & McCulloch, 2001). Both stages of antigenic variation and mechanisms known to effect *VSG* switching within each stage will be described in this introduction.

Antigenic variation is not unique to the trypanosome and occurs in, amongst others, pathogenic *Borrelia* species, *Plasmodium* spp., and the gut parasite *Giardia duodenalis*. Interestingly, the system has also been observed in parasites such as *Babesia* and once again *Plasmodium* spp. where the variant proteins carry functions additional to antibody evasion. In some ways, the trypanosome antigenic variation story serves as a model for understanding mechanisms allowing other microorganisms to evade immune systems and propagate in the host.

1.4 VSG coat structure

A dense VSG coat on the surface of the cell membrane encompasses both bloodstream and metacyclic form trypanosomes (Fig 2). This assembly appears to thwart complement activation (Ferrante and Allison, 1983), while preventing other non-humoral molecules from reaching the cell membrane, and shields the invariant surface molecules (which probably do not penetrate far into the coat) from immunoglobulin (Ig). Smaller molecules are capable of passing through the VSG coat and reaching membrane receptors. Receptors for larger macromolecules are found within the flagellar pocket (Grab *et al.*, 1992; Webster and Russel, 1993; Overath, 1994; Coppens *et al.*, 1998). This is readily accessible to large protein complexes and Ig but not the cellular arm of the host immune system (Borst, 1991). It is speculated that humoral immune responses to these invariant proteins may be partially redundant, in that antibodies may be cleared from the flagellar pocket by endocytosis (Borst and Fairlamb, 1998).

A "hinge" region separates the domains of each VSG monomer (Johnson and Cross, 1979; Carrington *et al.*, 1991), with each monomer containing 400-500 amino acids (Carrington *et al.*, 1991). Host antibody accessible epitopes are represented at the amino-terminal domain. This region displays a large degree of variation in amino acid sequence, with between 13 and 30% sequence identity between VSGs (Miller *et al.*, 1984; Carrington *et al.*, 1991). However, limited amino-terminus homology has been discovered between VSGs from the same serodeme (Olafson *et al.*, 1984) and also between immunologically distinct VSGs (Barbet, 1985). Following alignment of 19 VSGs (aligned on the basis of conserved cysteine residues), it was discovered that amino termini could be sub-classified into 3 distinct classes (types A, B and C) (Carrington *et al.*, 1991). X-ray crystallography has been used to deduce the three-dimensional structure of the variable domain of VSG molecules and they appear to fold into similar three-dimensional structures (Blum *et al.*, 1993). Carrington & Boothroyd (1996) have more recently reported that, despite different primary sequences, VSGs can perform the same apparent function of producing a monolayer barrier that can prevent host antibodies from recognizing other cell surface proteins due to similar tertiary structures.

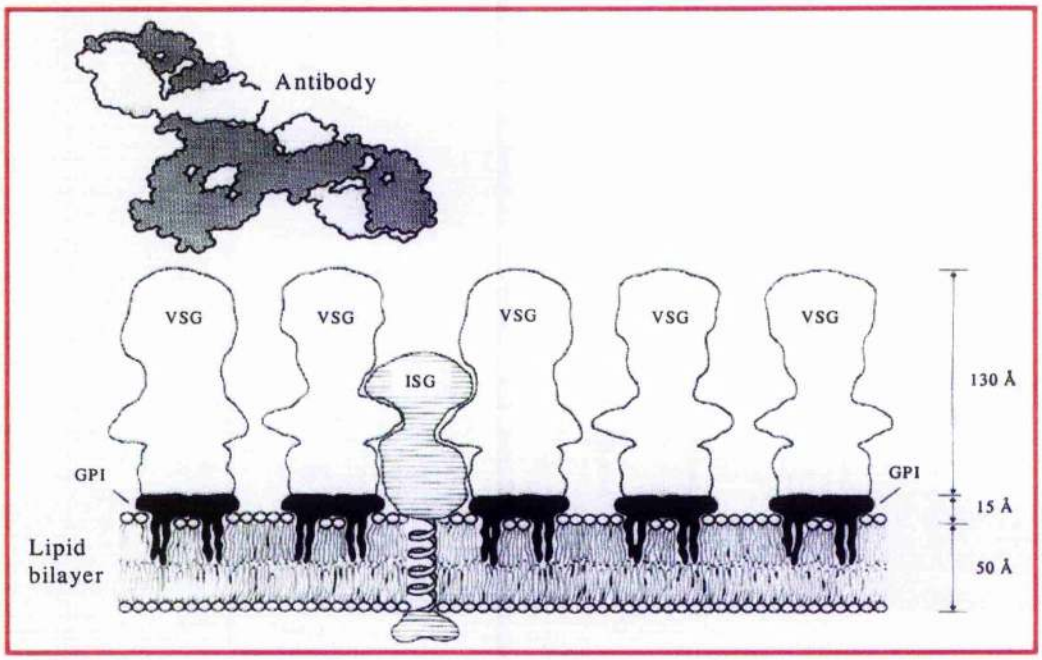

Figure 2. Dimeric membrane-form VSG molecules attached to the lipid membrane by GPI anchors. Invariant surface glycoprotein (ISG) and a host antibody are also represented. After Overath *et al.*, (1994).

The carboxy-terminus is more conserved, particularly in the 50 amino acids most proximal to the terminus (Allen and Gurnett, 1983; Carrington *et al.*, 1991). Cysteine residue alignments in this domain allow the classification of four distinct classes (types 1-4) (Carrington *et al.*, 1991). It was also found that amino-terminus classes A and B could be associated with each of the carboxy-terminal classes 1,2 and 3. A high degree of amino acid conservation occurs in VSGs across the carboxy-terminal hydrophobic tail of the VSG precursor peptide (Holder and Cross, 1981; Cross, 1990). The pre-polypeptide is modified in the endoplasmic reticulum, where the conserved tail is liberated from the VSG, and a glycophosphatidylinositol (GPI) moiety is covalently linked to the remaining protein. Post-attachment, these mature molecules are referred to as membrane form (mf) VSGs. During an antigenic switch the mature mfVSGs are exocytosed to the cell surface in the flagellar pocket and then diffuse laterally to form the new coat (Overath *et al.*,

1994), where the GPI anchor attaches the VSG to the phospholipid bilayer of the plasmid membrane.

Upon differentiation to the procyclic form, the VSG coat is replaced by a coat composed of the protein procyclin. The transition between life cycle stages is accompanied by the occurrence of a mixed coat, composed of both VSG and procyclin (Ziegelbauer and Overath, 1990; Roditi and Pearson, 1990). This may maintain a defence for the parasite prior to whole VSG shedding.

1.5 Bloodstream expression sites

Bloodstream form *VSG* genes are transcribed from sub-telomeric loci, termed bloodstream expression sites (*BES*s). These are large transcription units consisting of 40-60 kb of sequence (Fig 3a). They are polycistronic, containing a single *VSG* gene and a range of additional genes termed expression site associated genes or *ESAG*s (between 8 and 10) (Cully *et al.*, 1985; Pays *et al.*, 1989; Gottesdiener, 1994). The *VSG* gene is located directly upstream of the telomeric repeats. A battery of degenerate 70-bp repeats is located between the last of the *ESAG*s and the *VSG* gene (Liu *et al.*, 1983; Campbell *et al.*, 1984; Shah *et al.*, 1987). The common presence of these repeats upstream of internal and BES *VSG* genes suggests a conserved role in switching. The role of all *ESAG*s in the trypanosome life cycle has not yet been elucidated, though it is known that *ESAG6* and *ESAG7* encode the heterodimeric partners of a surface transferrin receptor (Schell *et al.*, 1991; Steverding *et al.*, 1995); *ESAG4* encodes adenylate cyclase (Paindavoine *et al.*, 1992; Alexandre *et al.*, 1996). *ESAG8* is apparently essential to the cell and its product accumulates in the nucleolus, although it is not required for regulation of the expression site (Hoek *et al.*, 2000; Hoek & Cross, 2001) and the recently discovered *ESAG11* appears to encode GPI-anchored glycoproteins (Redpath *et al.*, 2000). The bloodstream expression site (as well as the metacyclic expression site) is transcribed probably by RNA polymerase I, as is the case with ribosomal RNA (rRNA) and procyclin loci (Chung *et al.*, 1992; Lee and Van der Ploeg, 1997; Laufer *et al.*, 1999). Upstream of the *BES* lies a series of repeats, each unit 50 bp long. It is tempting to postulate that these repeats may

provide some form of DNA structure that is conducive towards attracting machineries involved in exacting recombination reactions. There are believed to be up to 32 bloodstream expression sites in the trypanosome genome, though this has yet to be confirmed (Vanhamme *et al.*, 2000).

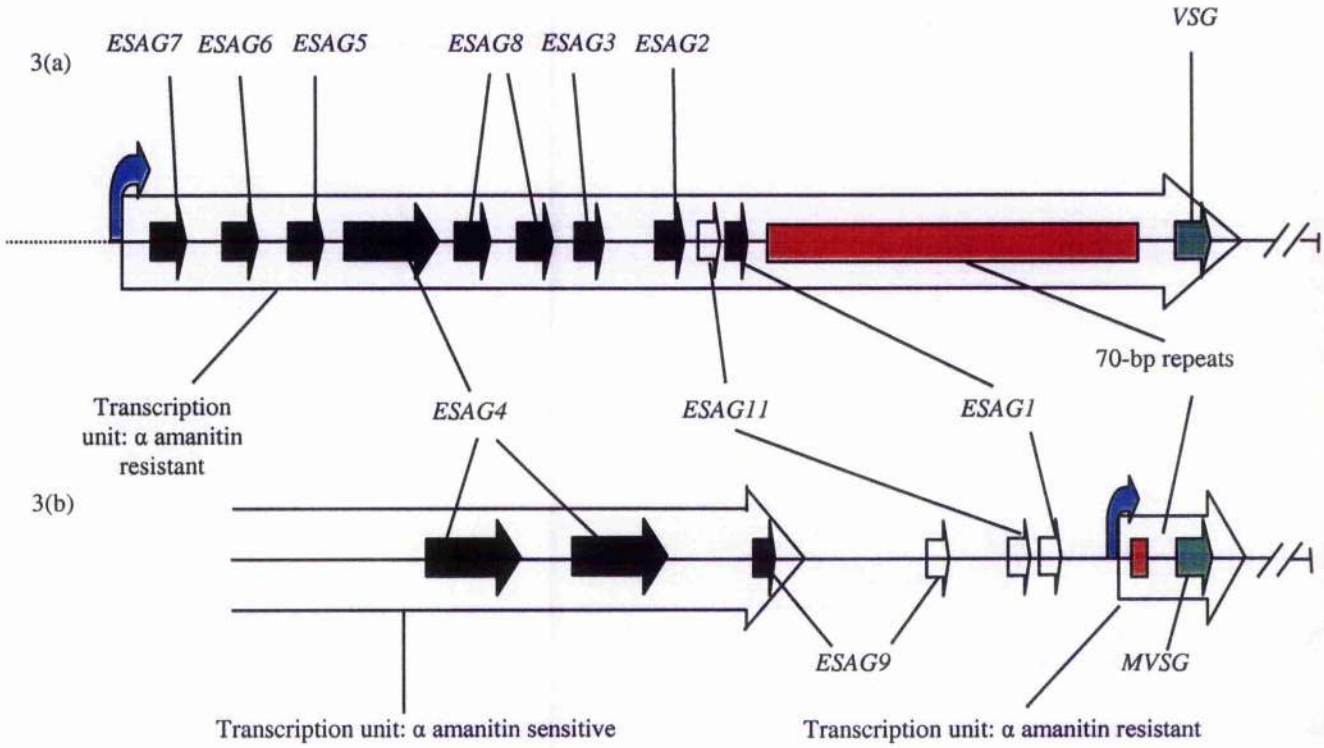

Figure 3. *VSG* expression sites in trypanosomes. (a) represents a telomere harbouring a bloodstream expression site, with the chromosome ending to the right. Genes are shown as black arrows and pseudogenes as empty arrows. The 70-bp repeats are shown in red. (b) shows a metacyclic expression site. Large empty arrows show transcription units for each expression site type and promoters are depicted as curved arrows. These diagrams are adaptations of those shown in Barry & McCulloch, 2001.

1.6 The genetic mechanisms for antigenic variation in the mammalian host

Two main switching mechanisms are utilised by the parasite to bring about a switch: transcription-based or recombination-based. Another class of switch mechanism has been proposed whereby rapid generation of point mutations occurs during the generation of a copy of a *VSG* gene (Donelson, 1995). It is not known, however, whether the occurrence of point mutations in *VSG*s would be sufficient to avoid specific immune responses previously elicited against the unaltered *VSG*. It has also been proposed that the experiments used to analyse these events were actually conducive to the selection of point mutations (discussed by Graham and Barry, 1996). For this reason this section is restricted only to the transcription- and recombination-based switching mechanisms.

1. Transcriptional *VSG* switching

On its own, this form of switching is limited by the number of expression sites in the trypanosome genome. Indeed, this represents a small number of the total number of *VSG*s in the genome. Transcriptional (*in situ*) switching results in the activation of an expression site concomitant with inactivation of the previously active expression site. Recent experiments have focused on elucidating whether activation and inactivation are linked and whether both newly and previously activated expression sites are completely on or off. This work has involved transgenic manipulation of trypanosomes harbouring known expression sites, by tagging *ES*s with antibiotic resistance markers (Zomerdijk *et al.*, 1990, 1991; Navarro and Cross, 1996; Horn and Cross, 1997; Rudenko *et al.*, 1998). It has ruled out DNA rearrangements as a means of regulating *ES* activity. Recent work, however, has demonstrated the presence of a novel nuclear body that correlates with transcription of an expression site (Navarro and Gull, 2001). This body co-localises with RNA PolI but not with other expression sites. Whether this structure is controlled stochastically or in a more regulated manner is not known. Neither is it known what factors are involved in the localisation of the expression site to the novel body.

2. Recombinational switching

DNA recombination is the most common route for *VSG* switching allowing a means of transferring *VSGs* (basic copy – BC) from transcriptionally silent areas to expression sites. Indeed, it has recently been confirmed that recombination reactions are important in dictating which *VSG* is expressed at any one time (Robinson *et al.*, 1999). Duplicative transposition of an internal *VSG* to an expression site (Expression linked copy – ELC) is a common strategy for activation of internal coat genes (Fig 4). The previously active *VSG* is deleted upon transposition of the new gene. These reactions are processed using homologous recombination, with substrates for recombination existing both upstream and downstream of the *VSG* gene.

The 5' flank of each BC *VSG* consists of several of the 70-bp repeat elements that are found in abundance in bloodstream expression sites. Homology with the *BES* also exists in the carboxy terminus-encoding region of each *VSG* gene as well as in its 3' untranslated region (UTR) (Timmers *et al.*, 1987; Lee and Van der Ploeg, 1987). The homology can extend up to 200 bp into the 3' UTR-encoding region, which contains 16 nucleotides conserved in most *VSG* genes (Majumder *et al.*, 1981). Interestingly, although the 70-bp repeats have been demonstrated to be involved in some *VSG* gene recombination (Florent *et al.*, 1987; Lee and Van der Ploeg, 1987), their absence does not affect the incidence of *VSG* duplications to the bloodstream expression site (McCulloch *et al.*, 1997), at least in the S427 strain. The downstream conserved regions have also been shown to act as targets for *VSG* recombination reactions (Pays *et al.*, 1983a; Timmers *et al.*, 1987). Most work carried out to date has involved low-frequency switching monomorphic cell lines, though it has been demonstrated that the 70-bp repeats delimit duplication boundaries in pleomorphic switchers (Delauw *et al.*, 1987; Shah *et al.*, 1987; Matthews *et al.*, 1990). Further work must be carried out, however, to delineate wholly the role of these repeat structures in switching in pleomorphic cells (pleomorphic trypanosomes maintain the ability to develop through all stages of the life cycle).

Another form of recombination involved in VSG switching is a mechanism referred to as telomere conversion. This strategy involves the same principles as internal gene transposition; with duplication of a telomeric *VSG* gene. The conversion limits, however, can fall beyond the boundary of the expression site, resulting in the replacement of large segments of sequence, and possibly the whole telomere (deLange *et al.*, 1983). This is believed to be the main route for activation of minichromosomal *VSG* genes (Barry *et al.*, 1997).

An additional recombination mechanism in the *VSG* system is reciprocal telomere recombination. Simply, the reaction involves the exchange of one telomeric region (including its *VSG*) for another (Pays *et al.*, 1985). This is not a commonly observed phenomenon, with frequencies matching the frequency of background cross-over events. However, recent reports suggest that what were previously described as transcriptional switches were actually telomeric recombination reactions (Rudenko *et al.*, 1996). This reaction has not yet been demonstrated in pleomorphic trypanosomes, probably because relatively it is very rare.

Another process used in *VSG* switching is that of mosaic gene formation. This results in the formation of a *VSG* gene hybrid and the production of a novel surface coat which permits evasion from existing antibodies in the bloodstream (Longacre and Eisen, 1986; Roth *et al.*, 1989; Thon *et al.*, 1989; Kamper and Barbet, 1992). These gene formations are seen to arise in chronic infections when the host has built up immunity to many non-mosaic genes (Roth *et al.*, 1989; Thon *et al.*, 1990; Barbet and Kamper, 1993). In its simplest form, the hybrid arises when a new *VSG* gene does not completely replace the previous sequence in the *BES*, producing a chimera composed mainly of the new gene (Pays *et al.*, 1983b) and encoding epitopes from the new gene. This event requires that the two genes harbour significant homology with each other outside the epitope-encoding region. A number of cases have been reported where mosaic genes are composed from various *VSG* members that are incomplete pseudogenes and cannot individually encode VSG coats (Thon *et al.*, 1989).

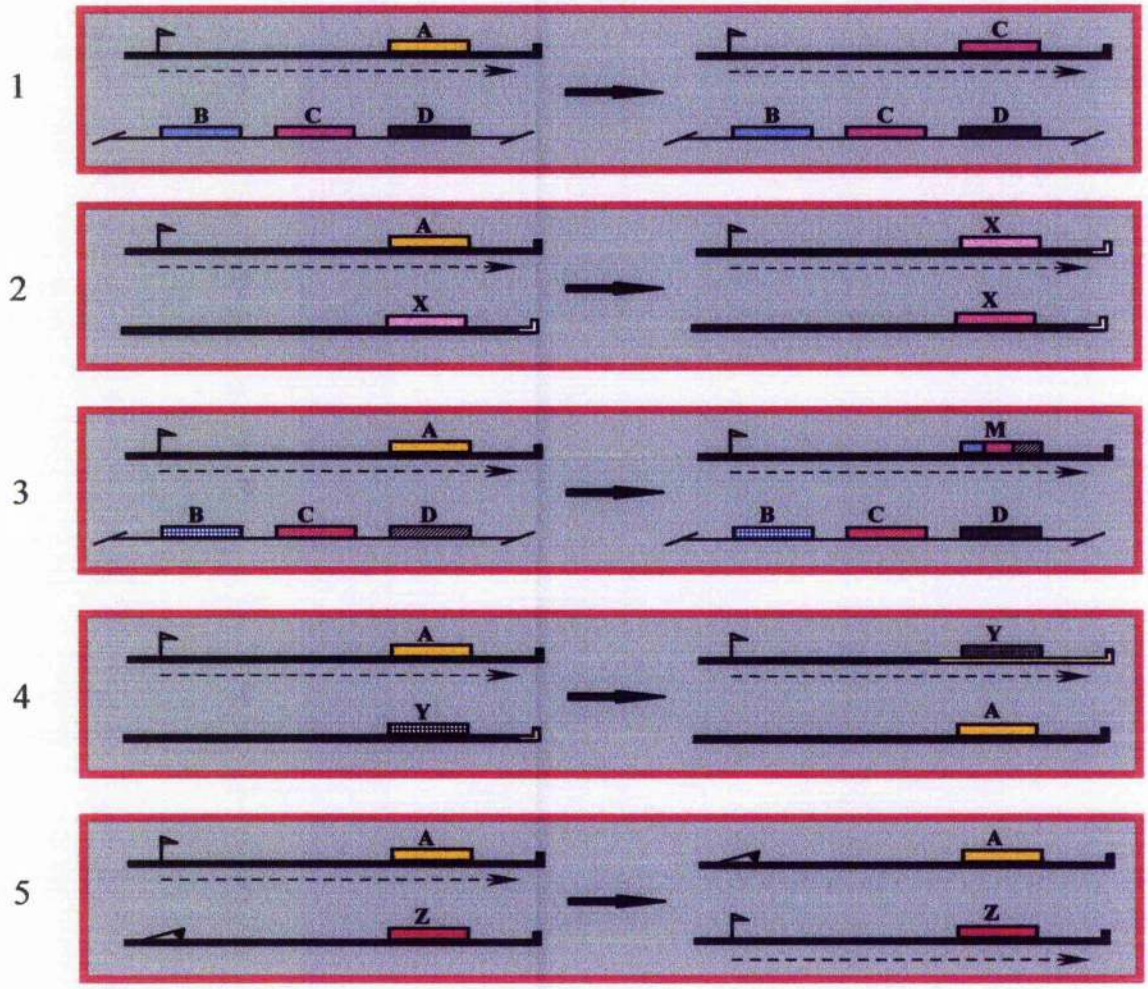

Figure 4. *VSG* switching in *Trypanosoma brucei* adapted from (Borst, 1991): (1) duplicative transposition of a chromosome internal *VSG* gene; (2) telomere conversion; (3) segmental gene conversion; (4) reciprocal telomere exchange; (5) "*in situ*" switching. Region A is the *VSG* gene in the active *BES*; B, C and D represent chromosome internal *VSG* genes; X and Y represent telomeric silent *VSG* genes; M represents a mosaic gene; Z represents a *VSG* gene present in a different *BES*. Expression site promoters are denoted by flags, which are raised or lowered to denote activation / silencing respectively; transcription is indicated by an arrow under the *BES*. The vertical bars downstream of the *VSG*s represent the telomere ends.

1.7 Differential expression of *VSG*s in metacyclic stage trypanosomes

The first appearance of a VSG coat in the *T. brucei* life cycle occurs in the salivary gland of the tsetse fly. This is the metacyclic stage of the trypanosome, which is preadapted for transmission to the mammalian host. Metacyclic *VSG*s are expressed from metacyclic *VSG* expression sites (*MES*s). These are a lot shorter than their bloodstream counterparts, from which they are believed to have evolved (Barry *et al.*, 1998) (Fig 3b). Characterised *MVSG* expression sites are 4-6 kb in length. There are up to 27 *MVSG*s in the metacyclic population, which might mean there are up to 27 ES located exclusively in sub-telomeric loci (Turner *et al.*, 1988). Several such expression sites have been characterised (Barry *et al.*, 1998; Pedram and Donelson, 1999; Alarcon *et al.*, 1999). Another difference from bloodstream expression sites is that MESs are monocistronic.

Both types of expression site are transcribed probably by RNA polymerase I (Chung *et al.*, 1992; Lee and Van der Ploeg, 1997). The only other known trypanosome monocistrons are transcribed by RNA polymerase III. Regions proximal to MESs have also been characterised and are believed to be composed of mainly non-coding, haploid sequence. The close proximity of MESs to telomeres has led to much speculation regarding a locus-specific silencing effect. As with bloodstream *VSG* expression, *MVSG* expression is believed to be exclusive (reviewed in Barry & McCulloch, 2001). This means that only one *MVSG* is active at any one time with the other expression sites remaining silent. As mentioned above, all characterised *MVSG*s are sub-telomeric and consequently reside within their own transcription unit. Therefore, for a *MVSG* switch to occur, an *in situ* activation is required. However, *MVSG* switches do not occur within the tsetse fly. Upon injection into the host organism, metacyclic *VSG* genes continue to be expressed for 5-7 days, after which the dominant switching reactions result in expression of bloodstream *VSG*s.

1.8 The DNA dependent protein kinase complex

Ku is a heterodimer formed from the two subunits, Ku70 and Ku80, and the name Ku is derived from the name of the patient from whom the antigen was originally isolated. Human Ku70 is 69 kDa and Ku80 is 83 kDa. These proteins are believed to have arisen from duplication and divergence of a common ancestor, which presumably functioned as a homodimer (Dynan & Yoo, 1998). The two Ku subunits associate tightly with one another, and they may form a tetramer when bound to the two DNA ends (Cary *et al.*, 1997). Together with the DNA dependent protein kinase catalytic subunit (DNA-PK_{cs}) and XRCC4, they form what is known as the DNA dependent protein kinase complex (DNA-PK). The complex is central to non-homologous end joining (NHEJ), the process in which dissimilar DNA ends can be rejoined with minimal error, preventing the lethal effects of double strand breaks (DSB) in the genome. *KU* homologues have been identified in yeast, which however appears not to have a functional DNA-PK_{cs} homologue.

The Ku complex is found abundantly in the nucleus, usually at telomeres, with translocation of Ku and other telomere binding proteins (Sir etc.) to new DNA breaks in *S. cerevisiae* (Martin *et al.*, 1999). In humans DNA-PK_{cs}, in association with DNA-bound Ku, mediates the catalytic function of DNA-PK. DNA-PK_{cs} can also bind to, and be activated by, DNA ends in the absence of Ku, but Ku is likely to be required for stabilisation of DNA binding by DNA-PK_{cs} *in vivo* (Critchlow and Jackson, 1998). It is widely believed that DNA-PK_{cs} and Ku provide a bridge-like structure that will facilitate subsequent re-ligation of two broken DNA molecules. Indeed, atomic force microscopy has shown that DNA-bound Ku is capable of juxtaposing DNA ends (Cary *et al.*, 1997). More recently the crystal structure of the human Ku heterodimers was determined both alone and bound to a 55-nucleotide DNA element (Walker *et al.*, 2001). Ku70 and Ku80 have a common topology and form a dyad-symmetrical molecule with a preformed ring that encircles duplex DNA. Ku makes no contact with the sugar-phosphate backbone, but it fits sterically to major and minor groove contours so as to position the DNA helix in a defined path through the protein ring. This structure seems well equipped to support broken DNA molecules and is sequence-independent, in keeping with its NHEJ role. Ku

is a multi-faceted protein and is a prime candidate for catalysis of *VSG* switching. The following sections describe the roles of Ku.

1.9 Ku and telomere position effect

Telomere Position Effect (TPE) is a gene-repression mechanism whereby chromatin influences the transcriptional status of genes in telomeric regions. Silencing is best understood in *S. cerevisiae* and *S. pombe*. Telomeric domains are not unique in their ability to cause this type of phenomenon, with silent donor mating-type loci and centromeres also being affected (reviewed by Sherman and Pillus, 1997; Klar *et al.*, 1998). The Ku complex effects a telomeric silencing phenomenon in yeast, suggesting a role in maintaining telomeric expression sites transcriptionally silent. *S. cerevisiae ku* null mutants are incapable of silencing genes harboured at sub-telomeric loci (Boulton and Jackson, 1996; Boulton and Jackson, 1998). This result suggested the Ku heterodimer as a candidate for causing the exclusive expression of *MVSG*s from expression sites, because their promoters lie within 5 kb of the telomere tract.

Telomeric silencing caused by the Ku heterodimer, along with a host of additional proteins, is believed to be due to the formation of a physical barrier of protein molecules preventing entrance of transcription factors (Fig 5). Interestingly, the TPE is not caused by Ku in *S. pombe* (Manolis *et al.*, 2001). Such observations were made in manipulated systems in *S. cerevisiae* whereby genes are artificially introduced into sub-telomeric loci. The trypanosome serves as an ideal model for testing the universal importance of telomeric silencing in dictating the transcriptional status of naturally telomeric and sub-telomeric loci. More importantly, genes involved in enhanced phenotypic variation inhabit these loci in trypanosomes. These genes are known as contingency genes and play, characteristically, an important role in the maintenance of pathogen populations within their hosts. Pathogenic microorganisms, including protozoa and fungi, commonly use contingency gene systems that are capable of creating phenotypic diversity during infection and can enable evasion of host immune systems (Moxon *et al.*, 1994; Deitsch *et al.*, 1997). A single contingency gene is expressed at any one time and this is usually part

of a large multi-gene family. It is common for members of such multi-gene families, including the expressed *VSG*, to be located telomerically. The functional importance of this location is not fully understood, repetitive sequence in the form of the telomeric repeats may be important substrates for recombination reactions that may be involved in switching the member being expressed. Alternatively, it has also been postulated that the telomere plays a more direct role in phenotype switching, by using effects such as TPE to facilitate switching between members of the *VSG* family (Horn & Cross, 1995; Rudenko *et al.*, 1995; Barry *et al.*, 1998; Scherf *et al.*, 2001).

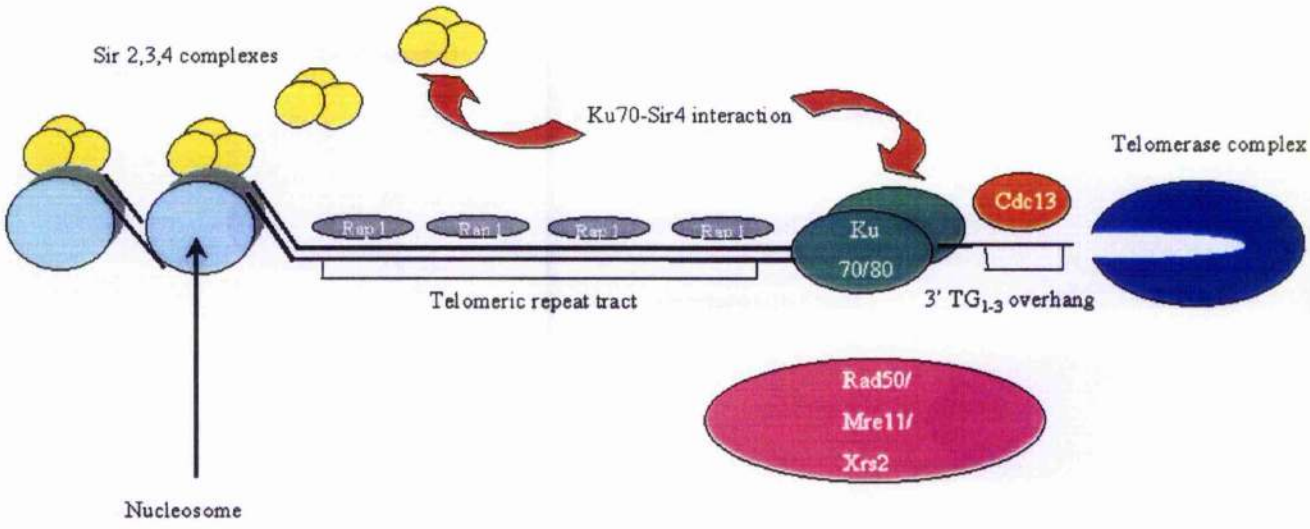

Figure 5. Architecture of telomeres in *Saccharomyces cerevisiae*. This cartoon depicts the best characterised telomere binding proteins in budding yeast. The telomere repeat tract is represented by two parallel black lines, while overhanging 3' sequence is shown as a single black line. RapI is the DNA binding element that allows the Sir complex to localise to telomeric sequence.

1.10 Ku and telomere length maintenance (TLM)

Ku also plays an important role in maintaining telomere tract lengths in yeast (Boulton & Jackson, 1998; Baumann & Cech, 2000). This phenotype is not observed in mammalian cells (Samper *et al.*, 2000), where, in fact, there is a slight increase in the length of telomeres when *KU80* is disrupted. Indeed, there are believed to be at least 25 proteins involved in this process, including the Sir (Sir2, Sir3 and Sir4) complex and Rap1 (reviewed by Shore, 1997; Diede and Gottschling, 1999). Ironically, many of the factors believed to be involved in this process are also pivotal in the repair of DNA breaks by NHEJ (see below). Telomeres are characteristically DNA-protein complexes where the interacting DNA is actually terminal.

This poses the quandary of why the absence of Ku (involved in NHEJ) causes chromosomes to fuse. Recent evidence has suggested, however, that a unique DNA structure exists at these termini that may prevent such fusions. This structure, known as a t-loop, is not detectable in yeast (Grunstein, 1997; de Bruin *et al.*, 2000, 2001) although a small DNA fold-back exists that could still provide a feasible reason for the lack of interchromosomal telomere fusion. t-loops have been demonstrated formally in *Oxytricha nova* (Murti and Prescott, 1999) and more recently, at the telomeres of minichromosomes in *T. brucei* (Munoz-Jordan *et al.*, 2001). These structures are believed to be formed by invasion of the characteristic 3' DNA overhang present at telomere termini into sub-telomeric duplex DNA (reviewed in Scherf *et al.*, 2001). Hence, this may represent a universal telomere structure. Recent evidence has suggested that native telomeric sequences are not necessary for the formation of t-loops *in vitro* (Stansel *et al.*, 2001). It is important, however, that the overhang sequence is 3' rather than 5' or blunt, the latter structures resulting in diminished t-loop formation.

t-loop formation and maintenance in humans is catalysed by TRF2, although it is not known how this reaction is carried out as TRF2 does not bind single-stranded telomeric DNA (Broccoli *et al.*, 1997). *In vivo*, it is likely that t-loop formation involves the participation of other proteins. The Mre11 complex, recruited onto telomeres by TRF2 (Zhu *et al.*, 2000), has been proposed to play a role in loop formation. It is likely that *in*

vivo, these structures help disguise telomeres from being recognised by native DNA repair machineries, preventing individual cells from entering cell cycle arrest. Additionally, the formation of such a structure is likely to inhibit physical interactions between naked telomeric DNA substrates and nucleases. This subsequently would allow telomere length homeostasis within limits and protection of sub-telomeric genes.

1.11 Ku and telomere tethering

A further telomere-related function of the Ku heterodimer in yeast involves tethering of telomeric sequences of *S. cerevisiae* chromosomes to the nuclear periphery (Fig 6) (Galy *et al.*, 2000). Previous reports demonstrated that disruption of Ku resulted in mislocalisation of telomeres in the yeast nucleus (Laroche *et al.*, 1998). Galy and colleagues realised the importance of Ku in docking reactions after they disrupted genes involved in nucleoporin complexes (NPCs). These mutant lines were defective in DNA DSB repair and concomitantly displayed alleviation of telomeric transcriptional repression. Indirect immunofluoresence studies with Ku fusion proteins showed that mutagenesis of a NPC protein, Mlp2, resulted in dispersal of Ku70 in the nucleoplasm. These results imply a role for NPCs beyond mediating nuclear-translocation processes, with the possibility that telomere localisation and chromatin organisation are also under the control of porin related proteins. Indeed, Andrulis *et al.*, (1998) demonstrated that perinuclear localisation of telomeres in yeast results in the establishment of transcriptionally silent chromatin.

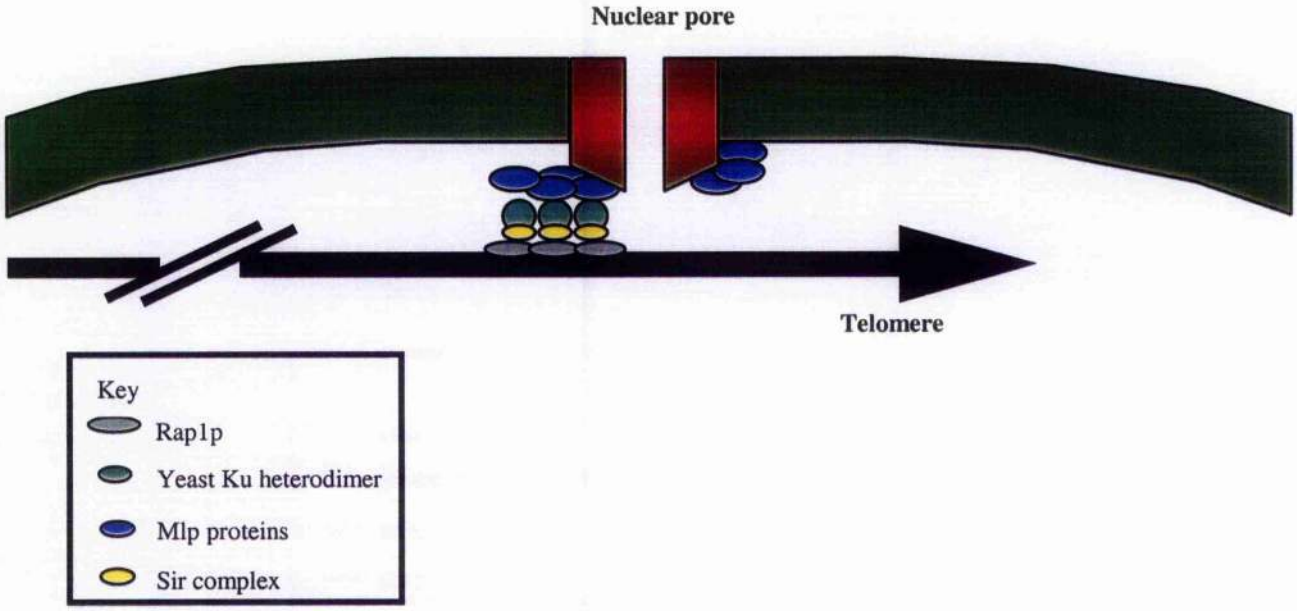

Figure 6. Telomere tethering to nuclear periphery in *Saccharomyces cerevisiae*. This cartoon represents protein complexes involved in telomere localisation in yeast. Telomeric DNA sequence is shown as a black arrow with chromosome internal sequence upstream of the parallel lines.

1.12 Ku and DNA break repair by Non-homologous end joining (NHEJ)

Processing of breaks by Ku and other complexes results in illegimate recombinational repair (see below). DNA ligation is the critical enzymatic step in NHEJ (Fig 7). The two broken DNA ends of a molecule are directly joined with no overlap (end-to-end) or with minimal overlap and the use of short fortuitous homologies near the two ends. "Simple ligation" i.e. the joining of two ends with cohesive protruding single strands or two blunt ends, a process that conceivably can be achieved by a DNA ligase alone, must also be classified as a form of NHEJ.

The model organism for study of this relatively simple form of DNA repair is *S. cerevisiae*, which possesses two ATP-dependent ligases that are homologues of the mammalian DNA ligases I and IV. The yeast DNA ligase IV homologue (Lig4) was

demonstrated to function epistatically with XRCC4 in NHEJ (Grawunder *et al.*, 1997). In yeast, homologous recombination is the predominantly observed mechanism for repairing double strand DNA breaks, but NHEJ is detectable when homologous recombination is inactivated e.g. by a mutation in *RAD52* (reviewed in Critchlow & Jackson, 1998; Taylor & Lehmann, 1998).

Other proteins are now known to contribute to NHEJ. The Mre11 complex is thought to process DNA ends after bridging reactions caused by Ku create accessible DNA. In mammalian cells, the Mre11 complex is composed of Mre11, Nbs2 and Rad50. The complex exhibits 3'- to 5'- exonuclease activity. Deficiencies in Nbs1 results in radiation sensitivity, chromosomal instability and the cancer predisposition condition Nijmegen breakage syndrome (NBS). It is believed that Nbs1 plays a role in DNA-damage signalling (van der Burgt *et al.*, 1996). Proteins from the Mre11 complex are also involved in forms of homologous recombination and are capable of performing homologous recombination reactions utilising small stretches of homology tracts, termed microhomologies. This reaction is catalysed preferentially when DNA DSBs breaks cannot be repaired using its intact sister chromatid or homologous chromosome as a recombination repair template.

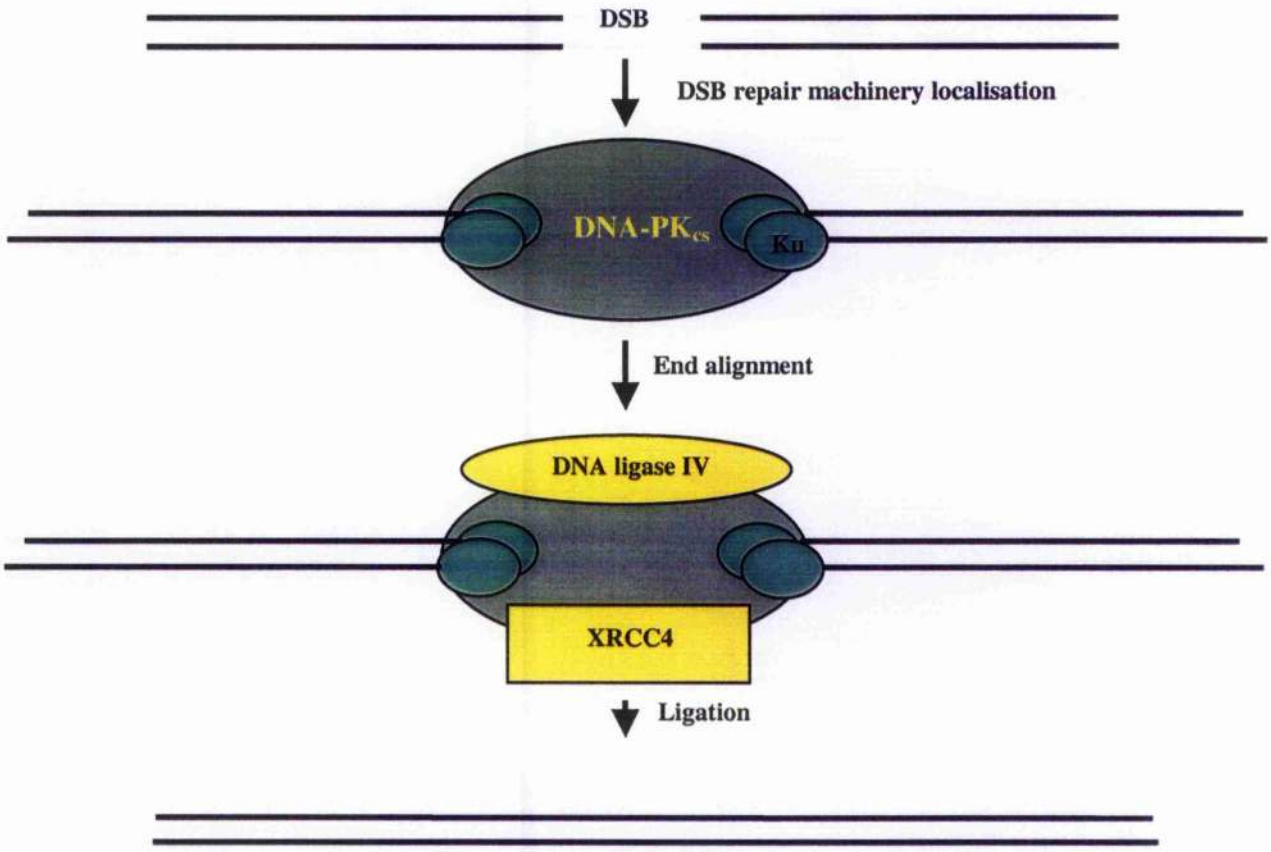

Figure 7. Double-stranded DNA repair by non-homologous end joining. Occurrence of a double-stranded DNA break is accompanied by binding to DNA termini by the Ku heterodimer and DNA-dependent protein kinase catalytic subunit (DNA-PK_{cs}). Following end alignment and resection, DNA ligase IV and XRCC4 rejoin the broken DNA molecule.

1.13 DNA repair by Homologous Recombination

Homologous recombination events require damaged DNA strands to base pair with a homologous partner (Fig 8), generally involving hundreds of perfectly matched base pairs, although shorter patches of homologous tracts can also be used. The genetic requirements for classical homologous recombination differ from those for illegitimate recombination, which occurs in only about one in a hundred cells when homologous repair is prevented in *S. cerevisiae*. Yeast *rad* mutants deficient in homologous

recombination, such as *rad51, rad52* and *rad54* mutants, are thus profoundly sensitive to ionising radiation, whereas illegitimate recombination mutants show a normal ability to repair radiation-damaged DNA. There are several types of homologous repair: single-strand annealing, gene conversion and break-induced replication. Each form of repair begins with 5' to 3' resection following exonuclease activity, leaving long, 3'-ended single-stranded DNA tails.

1. Single-strand annealing (SSA) (Fig 8a)

This is the simplest form of homologous recombination whereby resection exposes complementary regions of homologous sequences originally flanking the DSB, resulting in deletion of intervening sequences upon repair. SSA occurs with as little as 60 bp of homology, although it is much more efficient with 200-400 bp (Sugawara and Haber, 1992).

2. Gene conversion (Fig 8b)

This process competes with SSA. Resected ends of a double-strand break invade and copy sequences from a homologous template located on a sister chromatid, a homologous chromosome, or at an ectopic location. Mechanisms effecting these conversions have not yet been elucidated although it has been proposed that the original mechanism suggested by Szostak *et al.*, (1983) must be revised (Paques and Haber, 1999). The mechanism shown in Figure 8 is termed synthesis-dependent strand annealing (SDSA).

Although gene conversion is deemed as a less erroneous mechanism, SSA out-competes gene conversion in mitotic cells even when the donor and recipient are intrachromosomal (Fishman-Lobell *et al.*, 1992; Wu *et al.*, 1997), apparently because the resection of DNA ends is not restrained.

3. Break-induced replication (Fig 8c)

Break-induced replication (BIR) occurs when only one end of a DSB is able to engage in homologous recombination, for example in haploid or hemizygous chromosomes of G1 diploids, where there is no homologous chromosome. Sequences to the centromere-

proximal side of the DSB might be able to find homologous sequences elsewhere in the genome and create a nonreciprocal translocation by a process known as recombination-dependent DNA replication or break-induced replication (Fig 8c). This appears to occur when chromosomes lack telomerase, the enzyme that maintains repeated sequences at telomeres that protect ends from fusions and other types of recombination (Le *et al.*, 1999). Even when there is homology on both sides of a DSB, BIR (using only one of the ends to initiate recombination) appears to be in competition with gene conversion (Voelkel-Meiman and Roeder, 1990). This mechanism also accounts for the extensive DNA replication found in many gene targeting events (Morrow *et al.*, 1997).

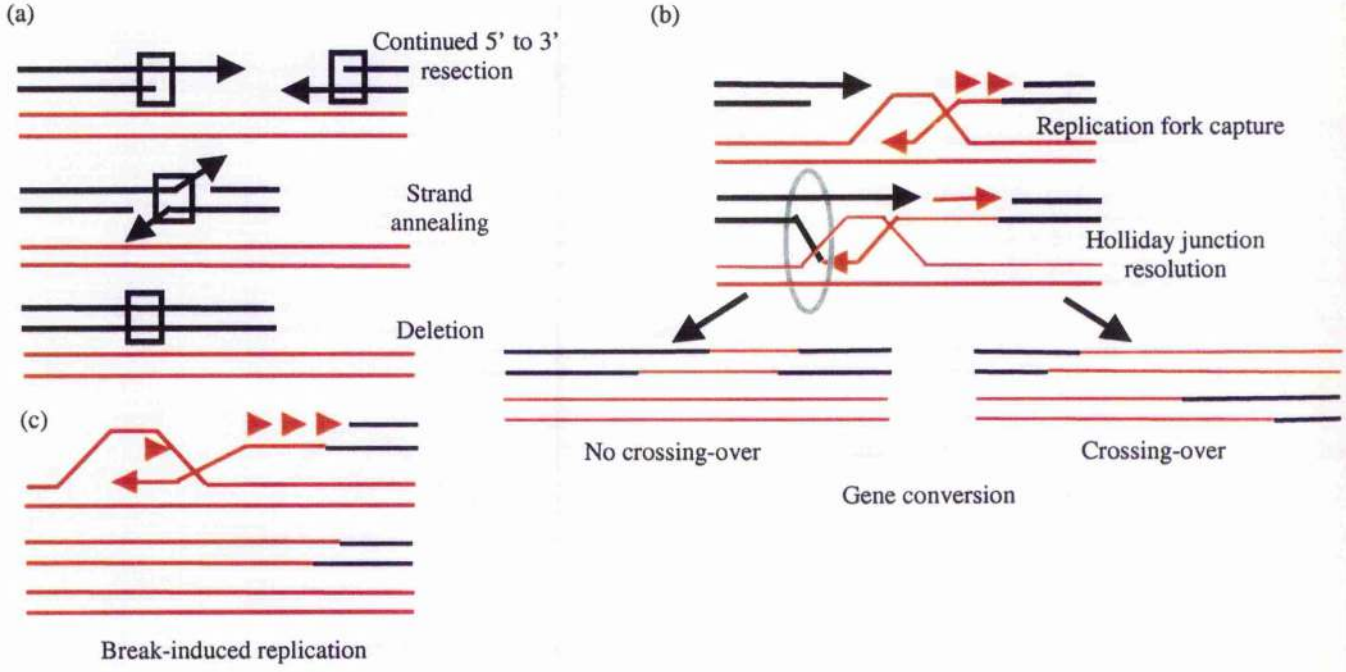

Figure 8. Homologous recombination. Following DSB occurrence resection occurs to leave a 3' overhang, which is capable of invading an intact template. Strand invasion is believed to introduce a modified replication fork, where both leading and lagging-strand DNA synthesis occurs. Branch migration then displaces newly synthesized strands. (a) If resection proceeds far enough to expose complementary strands of homologous sequences flanking a DSB (black boxes), repair can occur by single-strand annealing (SSA), leading to a deletion of all intervening sequence. (b) If the replication fork encounters the other end of the DSB, an intermediate containing two Holliday junctions can be formed, allowing gene conversions to be resolved both with and without crossing-over. (c) If the second end of the DSB fails to engage, replication can proceed all the way to the end of the chromosome (or until it encounters a converging replication fork). This process is known as break-induced replication (BIR).

1.14 *RAD* genes involved in homologous recombination

Genes involved in DNA repair were identified primarily as mutations sensitive to X-rays and subsequent genetic analysis allowed their classification as the *RAD52* epistasis group. A mutation in *RAD52* is as radiation sensitive as a double mutant with the other *rad* mutations placing *RAD52* as a gatekeeper of the molecular repair cascade. *RAD52* in *S. cerevisiae* stands alone in its pivotal role as the one gene that is required for all homologous recombination events. *RAD51, RAD54, RAD55* and *RAD57* have common phenotypes, being required for only some homologous recombination events. *RAD50, MRE11* and *XRS2* form another family of interacting proteins within this epistasis group whose deletions also have common phenotypes. This particular family is involved in both classical homologous recombination and illegitimate recombination. Homologues of the majority of these proteins have been identified in mammalian cells (Paques and Haber, 1999). Molecular cross-talk between DNA repair pathways (see previous section) may generate flexibility, allowing HR to repair a broad spectrum of lesions using different combinations of the RAD52-group proteins.

The eukaryotic homologue demonstrated to be involved in DNA repair in African trypanosomes is *RAD51* (McCulloch & Barry, 1999). Rad51 is a eukaryotic homologue of bacterial RecA with structural similarity, and is part of the Rad52 epistasis group. It has also been demonstrated to have a role in DSB repair (mitotic and meiotic) in *S. cerevisiae*. Like *E. coli* RecA, Rad51 can catalyse an ATP-dependent strand exchange between a single-stranded circular molecule and a homologous linear duplex (Sung, 1994; Sung & Robberson, 1995; Sung & Stratton, 1996; Namsaraev & Berg, 1997). Also, Rad51p seems to bind dsDNA much more strongly than does its bacterial counterpart (Benson *et al.*, 1998; Sung, 1994). There is debate regarding the polarity of the strand exchange reaction catalysed by Rad51p, with evidence beginning to show that Rad51 can deal with either the 3' or 5' DNA end (Namsaraev & Berg, 1997).

Little is known regarding factors causing DNA repair in trypanosomes. Recent studies have shown, however, that the trypanosome homologue of *RAD51* is involved in DNA repair of double-stranded breaks introduced to the trypanosome genome via the DNA

damaging agents methylmethane sulfonate and 3-aminobenzamide (McCulloch and Barry, 1999). DNA breaks introduced in this manner are generally repaired by homologous recombination. Disruption of this gene also resulted in reduced efficiencies of *VSG* switching, suggesting that switching is, at least in part, catalysed by Rad51 and involves homologous recombination, or alternatively Rad51 has a regulatory role in switching. This was the first demonstration that disruption of a specific gene in trypanosomes has an effect on antigenic variation. *rad51*$^{-/-}$ cells permit detection of alternative forms of DNA repair in African trypanosomes.

Work presented in this thesis tackled questions addressing both antigenic variation and DNA repair in African trypanosomes. The primary objective of the work involved testing the hypothesis that the trypanosome has orthologues of *KU70* and *KU80* that are involved in *VSG* switching and differential *VSG* expression. Follow-up experiments, looking for the occurrence of NHEJ, led to an analysis of DNA repair mechanisms.

CHAPTER 2

MATERIALS AND METHODS

Reagent Abbreviations

DAPI 4, 6-diamidino-2-phenylindole

DEPC diethyl pyrocarbonate:
Used at 0.1% to remove RNAase

EtBr ethidium bromide

NDS solution for the manufacture of genomic plugs (1 x):
0.5 M EDTA, 0.01 M TRIS base, 1% lauroyl sarcosine, pH adjusted to 8 or 9

PBS phosphate buffered saline (Sigma, Ltd.)

PSG phosphate/ sodium chloride/ glucose buffer (1 x):
0.06 M Na_2HPO_4, 3.6 mM NaH_2PO_4, 46 mM NaCl, 55mM glucose, pH to 8

SDS sodium dodecyl sulphate

SSC sodium chloride/ sodium citrate solution (1 x):
0.15 M NaCl, 0.015 M $Na_3C_6H_5O_7$

TAE TRIS/ acetate/ EDTA buffer (1 x):
0.04 M TRIS base, 0.04 M glacial acetic acid, 1 mM EDTA

TBE TRIS/ borate/ EDTA buffer (1 x):
0.089 M TRIS base, 0.089 M ortho-boric acid, 2 mM EDTA

TBS TRIS buffered saline (Sigma, Ltd.)

TE 10 mM Tris.Cl, 1mM EDTA

2.1 Trypanosome strains and transformation

Bloodstream form cells, *T. brucei* 221a trypanosomes (MITat1.2a) of strain 427 (Cross, 1975), or the transgenic derivative 3174 (McCulloch et al, 1997) were used and grown *in vitro* at 37°C in HMI-9 medium (Hirumi and Hirumi, 1989). EATRO (East African Trypanosomiasis Research Organisation) 795 was the procyclic strain of choice and was maintained *in vitro* at 27°C in SDM-79 (Brun and Schonenberger, 1979). Bloodstream form trypanosomes were transformed by electroporating 5 x 10^7 cells in 0.5 ml of Zimmerman Post-Fusion medium (132 mM NaCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 0.5 mM Mg acetate, 0.09 mM Ca acetate, pH7.0) supplemented with 1% glucose, using a Biorad Gene Pulser II set at 1.4 kV and 25 μF capacitance. Procyclic form trypanosomes were transformed in an almost identical manner. On this occasion, however, 2.5 x 10^7 cells were electroporated and the Zimmerman Post-Fusion medium was not supplemented with glucose. For either electroporation ~5 μg of restriction enzyme digested, phenol/chloroform extracted and ethanol precipitated plasmid DNA was used. Cells were allowed to recover for 18 h before drug selection. Bloodstream form transformants were selected on semi-solid agarose plates, whereas procyclic transformants (10^5 electroporated cells) were selected in 10 ml of liquid medium supplemented with ~10^6 untransformed cells. Bloodstream form wild-type and *rad51-/-* mutant cells were recovered in a different way. This was done by diluting 10^7 wild-type cells and 2 x 10^7 *rad51-/-* cells in 18 ml of HMI-9 medium containing the appropriate antibiotic, and distributing them over 12 wells of a multi-well plate in 1.5 ml samples. This method of limiting dilution does not guarantee clonal selection, but is a more reproducible method for determining transformation efficiency (which was calculated by the numbers of wells that grew out after 7-14 days) than selection on semi-solid agarose. pCP101 transformants were selected with 10 μg.ml^{-1} blasticidin S (Calbiochem). Other antibiotics used for selection included puromycin (1.0 μg.ml^{-1}) (Calbiochem), hygromycin (5.0 μg.ml^{-1}) (Roche), bleomycin (2 μg.ml^{-1}) (Cayla) and G418 (2.5 μg.ml^{-1}) (Sigma).

2.2 Isolation of genomic DNA

2.2.1 Preparation of genomic plugs from live trypanosomes

Trypanosomes were spun down gently (1500*g*) at room temperature and carefully resuspended in PSG buffer. This wash was repeated three times, and the concentration of the cells was determined using a haemocytometer under a phase contrast microscope. The trypanosomes were then spun down again and resuspended in PSG to a concentration of 1 x 10^8 cells.ml^{-1} (twice the final concentration of the plug) and warmed to 37°C for 1 min. Equal volumes of the trypanosome suspension and a 1.4% low-melting point agarose (InCert agarose, FMC Bioproducts) solution (which had been melted and cooled to 37 °C previously) were mixed together by swirling, resulting in a final concentration of 5 x 10^7 in 0.7% agarose. 100 µl of this suspension was pipetted into each well of a plug mould (BIO-RAD) and allowed to set. The plugs were then placed in NDS at pH 9.0, supplemented with 1 mg.ml^{-1} proteinase K (Sigma, Ltd.), at 50°C for 24 h. After this period, the plugs were transferred to NDS at pH 8.0, and supplemented with 1 mg.ml^{-1} proteinase K, at 55°C for 24 h. The plugs were then stored in fresh NDS at pH 8.0 at 4 °C.

2.2.2 Genomic DNA preparations

5-10 x 10^7 trypanosomes were harvested by centrifugation, resuspended in 50mM Tris.Cl (pH 8.0), 1mM EDTA, 100mM NaCl and lysed by addition of SDS to 1% in the presence of 100 µg.ml^{-1} proteinase K overnight at 37°C. This was phenol/chloroform extracted, ethanol precipitated and DNA resuspended in 50 mM Tris.Cl (pH 8.0), 1 mM EDTA.

2.3 Generation and analysis of trypanosome *VSG* switchers

The method used to measure frequencies at which the various cell lines switch their VSG coat has been described previously (McCulloch and Barry 1999). Switching

frequencies were carried out for wild-type, *KU70* heterozygous and *ku70* homozygous mutant cell lines. Mice immune to the VSG221 coat were generated when a number of ICR mice were intraperitoneally injected with 2 x 10^5 to 4 x 10^5 3174 trypanosomes (McCulloch et al., 1997) growing on Hygromycin (5.0 $\mu g.ml^{-1}$) and G418 (2.5 $\mu g.ml^{-1}$); 3-4 days later, the infections were cured by injection of cymelarsan (Rhone Merieux; 5 $mg.kg^{-1}$). The immunized mice were used between 7 and 35 days post-curing. Switch variants were selected when *KU* transformants were passaged from medium containing hygromycin and G418 into nonselective medium at a density of 0.5 x 10^4 to 1.0 x 10^4 $cells.ml^{-1}$ and grown until they reached densities of 1.0 x 10^6 to 3 x 10^6.ml^{-1}. Measured numbers were then injected into immunized mice in 300 µl of HMI-9. Twenty four h later, the mice were bled by cardiac puncture and trypanosomes isolated and cloned as described previously (McCulloch *et al.*, 1997). In calculating the *VSG* switching frequencies from the numbers of clones that grew out, we have assumed (as before, McCulloch *et al.*, 1997) that the 3174 *KU* wild-type, 3174 $KU^{+/-}$ and 3174 $ku^{-/-}$ cell populations double every 8 h. The drug sensitivities of the switched variant clones, as well as the presence or absence of the marker genes within the marked *VSG221* expression site, were assayed as described previously (McCulloch *et al.*, 1997).

2.4 Gel electrophoresis and Southern blotting

2.4.1 Gel electrophoresis

DNA separations were performed on 0.7% agarose gels (GibcoBRL, Life Technologies) run at 100 V in 1 x TAE buffer, using a commercial 1 kb ladder as a size marker (GibcoBRL, Life Technologies). Separations on which Southern blots were carried out were performed on agarose gels containing 1 x TBE buffer.

2.4.2 Genomic DNA digestions

Genomic DNA (usually 1 µg) was digested for 4 h with the appropriate restriction enzymes, following the manufacturer's protocol (New England Biolabs or Promega).

The products were fractionated on a 0.6%, 0.7% or 1.0% agarose gel run at 30 V (overnight in 1 x TBE) to ensure a high resolution of the bands, after which the DNA was transferred to a Nylon membrane by Southern blotting (see below).

2.4.3 Pulsed field gel electrophoresis (PFGE)

Chromosome sized DNA was resolved by PFGE using the CHEF-DR III system (BIO-RAD). Buffer composition is extremely important for PFGE and the gel must be made from (and the genomic plugs dialysed in) the buffer that is circulating in the tank (and not from a fresh stock). At least 2 l of the appropriate buffer were circulated in the electrophoresis tank for 10 min, after which a small sample was taken for dialysing the genomic plugs. The plugs were dialysed, at room temperature, in 1 ml of buffer, and then were placed at 4°C and left overnight. The following day, 110 ml of buffer were taken from the tank and the agarose gel was prepared at the appropriate concentration (SeaKem Gold agarose (FMC BioProducts)). Following dialysis, the genomic plugs were placed on the comb of the gel former (one-half a genomic plug per lane) and 100 ml of the molten agarose was then poured, and allowed to set. After 30 min, the comb was removed (leaving the plugs embedded in the gel) and the wells were filled with the remaining liquid agarose. The PFGE conditions utilised in this thesis were as follows: (i) 6–day general separation, 1.2% agarose gel at 15 °C in 0.089M Tris-borate, 0.1mM EDTA (1 x TB(0.1)E) (85 V, 1400-700 s pulse time, 144 h) (ii) Analysis of minichromosomes: Electrophoresis was conducted at 12°C and 4.6 V cm^{-1} for 42 h using a 1% agarose gel. The included angle between electrode states was 120° and the electrode switching time was ramped linearly from 8 to 15 s.

2.4.4 Southern blotting

Prior to blotting, the agarose gels were stained with EtBr and viewed and photographed on a UV transilluminator. The gel was then placed in 200 ml 0.25 M HCl for 15 min (to nick the DNA), rinsed with distilled water, and immersed in 200 ml denaturation solution (0.5 M NaOH, 1.5 M NaCl). After 30 min of denaturing, the gel was rinsed briefly with

distilled water and then placed in 250 ml neutralising solution (1 M Tris-HCl pH 8.0, 1.5 M NaCl) for 30 min. The DNA was then transferred to a Nylon membrane (Micron Separations, Inc. or Hybond-N) by wet blotting using 20 x SSC as the transfer buffer (Sambrook *et al.*, 1989); standard blots were left for 24 h, while PFGE blots were left for 48 h. After transfer, the DNA was crosslinked to the membrane using a UV spectrolinker (Stratagene).

2.5 Cloning of DNA fragments

2.5.1 Ligations

Following fractionation on a 0.7% agarose gel, DNA fragments were gel purified using the QIAgen gel extraction kit (following the manufacturer's protocol) and were cloned into pBluescript (Stratagene), PCR Script Amp SK(+) (Stratagene) (following the manufacturer's protocol) or Topo vector (Invitrogen) (following the manufacturer's protocol); the ligations were performed at room temperature for 4 h for pBluescript. Ligations using the Topo vector and PCR Script Amp SK(+) were carried out in accordance to the manufacturer's protocol. At each cloning attempt several different insert:vector ratios were prepared (*e.g.* 1,3,5 and 7 μl of purified product) to maximize the probability of a successful ligation. The PCR Script system works by adding an infrequently cleaving restriction enzyme, which cuts at a site within the plasmid polylinker (producing blunt overhangs) into the ligation mix. PCR products generated by *Taq* polymerase must be "polished" (to remove the 5' A overhangs) using *Pfu* polymerase before the ligation step. Some restriction enzyme digest products required overhang blunting prior to ligation reactions. Before the blunting reaction could be carried out, the restriction enzyme(s) used to cleave the DNA fragment was (were) heat inactivated. 3' overhangs were blunted using Pfu polymerase (Stratagene) (following manufacturer's protocol), while 5' overhangs were blunted using T4 DNA polymerase (New England Biolabs) in the presence of 10 mM dNTPs.

2.5.2 Transformations and plasmid retrieval

Heat shock was used to aid transformation of competent *E. coli* Xl-1 blue MRF cells (GibcoBRL, Life Technologies). 4 µl of the ligation product was added to a 100 µl sample of the competent cells (previously thawed on ice), mixed gently, and left on ice for 30 min. Heat shock was then carried out at 42 °C for 65 s, and the cells were immediately transferred to ice for 1-2 min. Afterwards, 0.9 ml of L-broth was added to the cells, and the mixture was subsequently incubated at 37 °C for 1 h. The cells were then centrifuged, resuspended in 100 µl L-broth, and spread over L-agar plates containing 0.27 M ampicillin (Sigma, Ltd). Transformed colonies were re-streaked on fresh L-agar plates (also containing ampicillin), after which a single colony was selected to inoculate into 5 ml L-broth (supplemented with ampicillin to a final concentration of 10 $mg.ml^{-1}$). Plasmids were prepared from 3 ml of the overnight culture using the QIAGEN Mini kit (following the manufacturer's protocol).

2.5.3 Cloning and sequencing *T. brucei KU70* and *KU80* genes

Sanger database entries 0a2 and a27e10 were recognised as putative *KU70* and *KU80* homologue fragments respectively, following alignment with eukaryotic homologues. These regions were amplified using DNA specific oligonucleotides:

KU705'	CCGAATTCGGGGAGGGTACGTCAGCGGG
KU703'	CC**AAGCTT**CAAGCTCTCGATTTGGATAGC
KU805'	GAATTCTTCCTCTGCGTGCCGTTGAC
KU803'	GAATTCGCACCTCTTCACCTACCGTC

Table 1. Oligonucleotides to amplify *KU70* and *KU80* open reading frames.

(*Eco*RI underlined, *Hind*III in bold). PCR was carried out using *Taq* DNA polymerase with ~3 pmoles of each primer, an initial 10 min denaturation at 95°C followed by 30 cycles of 95°C for 1 min, 65°C for 1 min and 72°C for 1 min. These reactions were carried out in 1 x buffer IV with separate $MgCl_2$ (reagents supplied by Advanced

Biotechnologies) and deoxynucleotides (to a total final concentration of 400 μM (Amersham Pharmacia Biotech)). Products of 248 bp and 214 bp were cloned into pBluescript (Stratagene) for 0a2 and a27e10 respectively. Sequence was determined on both strands using ABI-PRISM automated sequencing with custom-designed primers. Insert sequence was subsequently used to probe an ILTat 1.2 genomic library (*Sau*3A partial digests in λ-GEM-12). This initially allowed isolation of two fragments for each gene believed to contain putative *KU70* and *KU80* sequence. Once again, these clones were sequenced on both strands by ABI-PRISM automated sequencing, yielding the open reading frames (ORFs) described in Results. The remainder of each gene was sequenced on both strands in a similar manner using the following oligonucleotides:
T3: AATTAACCCTCACTAAAGGG; T7: GTAATACGATCACTATAGGGC.

KU70R1	CCGCAACAGCGCACACAACC	**KU70T1**	GGGTAATCAGGATGTGTGCA
KU70R2	GCAATCTGCTGCGTTCAGCAA	**KU70T2**	GTATTGTACAGCACCAACGC
KU70R3	ACCTGCGGCTTCAACGCAAT	**KU70T3**	CCCGTGAACTGTTTGTTGC
KU70R4	GCCACCTGCACATATCAGTA	**KU70T4**	ATCTGGGTGGAGGACGCACG
KU70R5	TGCTGACGACGTTCGTACCA	**KU70T5**	TACGAAGGGAGCAATACCGC
KU70R6	TATCTGCCGTGGAGTCGGCGT	**KU70T6**	GTAATGTGGTAGTGAGTCC
		KU70T7	AAGGTTTTGGCGCATGGCC

Table 2. *KU70* sequencing oligonucleotides.

KU80S1	TTGCTGCTGGAGGCGGTCGG	**KU80U1**	ATTGCTGCTCGCGTAGGGAA
KU80S2	GTGCTGTTCTCTCGTTCTCT	**KU80U2**	AGGAGCAACAAGGTGTGGGA
KU80S3	AACCTTCACCTGTCACTTGG	**KU80U3**	TAGCTACGAACGTGCTCGGT
KU80S4	GAATTCGCACCTCTTCACCTA CCGTC	**KU80U4**	AACGACAGGTGACGAAGATG
KU80S5	GAAGGGCACATGGTCTTCAG		

Table 3 *KU80* sequencing oligonucleotides.

2.6 Construct generation

2.6.1 *T. brucei KU70* and *KU80* knockout constructs

Two constructs were used to delete each gene, using two rounds of transformation. Each construct contained approximately 400 bp of targeting sequence, derived from sequences immediately upstream and downstream of either ORF, meaning that transformants delete the entire ORF. The targeting flanks were separated by one of two antibiotic resistance cassettes. Targeting flanks for each gene were amplified by PCR of genomic DNA using *Pfu* DNA polymerase (Stratagene).

***KU70* 5' flank primer (A)**	5' CC*TCTAGA*GAGGCGCTTCT CAATCTAAT 3'	***KU80* 5' flank primer (A)**	5' CC*TCTAGA*CAGCCAGC AGCGCTTCAATA 3'
***KU70* 5' flank primer (B)**	5' CC*TCTAGA*TATCACGCGGT ACCCAGGTT 3'	***KU80* 5' flank primer (B)**	5' CC*TCTAGA*ACGGAAAG CCATATTGGAG 3'
***KU70* 3' flank primer (A)**	5' CCTTGGGCCCAACCATTG CGAGTGCGCGTT 3'	***KU80* 3' flank primer (A)**	5' CCTTGGGCCCGAGAA ACAAAGTCACATATAAAT A 3'
***KU70* 3' flank primer (B)**	5' CCTTGGGCCCAGAACTCG AGCCTTGCAGTGGGATTCTCT A 3'	***KU80* 3' flank primer (B)**	5' CCTTGGGCCCAGAACT CGAGTGAGATGTGGCGAA CAGAGGAGTA 3'

Table 4 *KU70* and *KU80* knockout construct oligonucleotides. *Apa*I restriction enzyme sites are underlined, *Xho*I sites are in bold and *Xba*I sites are italicised.

Each 5' flank product was digested with *Xba*I and they were cloned independently into both pTBT and pTPT (gift of M. Cross and P. Borst, the Netherlands, Cancer Institute). pTBT contains the 400 bp blasticidin S deaminase ORF (BSR) flanked by 240 bp of 5' and 330 bp of 3' processing signals from the 5' and 3' flanks of *T. brucei* α-tubulin. pTPT contains identical processing signals flanking the 600 bp puromycin N-acetyltransferase ORF (PAC). Amplified 3' targeting flanks for each gene were independently digested with *Apa*I and cloned into both pTBT and pTPT already containing respective 5' targeting flank. The final transformation constructs were digested with *Not*I and *Xho*I prior to being transformed into *T. brucei*, and transformants were selected with 10 $\mu g.ml^{-1}$ blasticidin or $1\mu g.ml^{-1}$ puromycin.

2.6.2 *KU70* re-expression construct and pCP101

Cassette pCC101 was used to re-introduce *KU70* to its original locus. The entire *KU70* ORF along with 620 bp of 5' flanking sequence was cloned into an *Hpa*I restriction enzyme site. This cassette also contained 400 bp of actin intergenic sequence immediately downstream of the *KU70* ORF. This, upon correct integration, would provide 5' splicing of marker (bleomycin phosphotransferase ORF) RNA. Tubulin intergenic sequence (330 bp) provided polyadenylation signals along with a means of targeting the originally disrupted *KU70* locus. The plasmid was linearised with *Not*I prior to electroporation and transformants were selected on semi-solid agarose plates containing 2.0 $\mu g.ml^{-1}$ bleomycin (Invitrogen).

pCP101 contains the Blasticidin S Deaminase ORF flanked upstream by 240 bp of $\beta\alpha$ tubulin intergenic region, and downstream by 400 bp of actin intergenic sequence; further upstream lies a ribosomal DNA promoter (510 bp). This was digested with *Xho*I and *Eco*RV prior to transformation.

2.7 Telomere Length Maintenance Assay

Each cell line under analysis was sub-cloned on semi-solid agarose HMI-9 plates. Trypanosome colonies were subsequently maintained in HMI-9 liquid medium. Genomic DNA was harvested and digested with *Eco*RI or *Age*I depending on the Expression Site under analysis. *Eco*RI cuts ~2.7 kb upstream of the telomere tract downstream of *VSG221*, which is located in the active Expression Site. *Age*I cuts within the *VO2* gene in the transcriptionally inactive Expression Site analysed. DNA digests were size-fractionated on a 0.6% agarose gel. Each gel was Southern blotted onto Hybond N (Amersham) and probed with the N-terminus encoding region of either *VO2* or *VSG221*. Once more the filters were washed, after hybridisation, to a final stringency of 0.2 x SSC, 0.1% SDS at 65°C.

2.8 Retrieval of *T. brucei* genomic DNA sequence flanking integrated construct

Sequence immediately surrounding integrated plasmid DNA was retrieved by 'Topo-walking' (Invitrogen), a procedure that uses *Vaccinia* virus topoisomerase I to join a specific DNA linker onto a DNA fragment generated by primer extension from a known sequence. Genomic DNA was digested with *Hae*II and dephosphorylated with Calf Intestinal Phosphatase prior to *Taq* polymerase primer extension with either BSDfor or BSDrev primer. The following gene specific oligonucleotides were then used with LinkAmp1 primer (Invitrogen).

LinkAmp1	AGGCACAGTCGAGGACTTATCCTA
BSDfor	ATAGAATTCATGGCCAAGCCTTTGTCTC
BSDrev	TTTGAATTCTTAGCCCTCCCACAC
InvBSD1	CTTGAGACAAAGGCTTGGCC
InvBSD4	TGGGATTCGTGAATTGCC

Table 5 Topo-walking oligonucleotides

These reactions used 10 min 95 °C denaturation, followed by 30 cycles of 1 min at 95°C, 1 min at 55 °C and 3 min at 72 °C and finally 10 min at 72 °C. Amplified products were cloned into the Topo cloning vector (Invitrogen) and sequenced by ABI-PRISM automated sequencing.

2.9 Reverse transcription polymerase chain reaction (RT-PCR)

2.9.1 RNA isolation

RNA was isolated from trypanosomes grown in culture (after a density of ~10^6 cells.ml^{-1} was reached) with TRIzol (Gibco-BRL, Life-Technologies). The cells were centrifuged (3000*g*) at room temperature and resuspended in ~100 μl HMI-9 medium. The suspension was mixed with TRIzol (0.1ml suspension: 1 ml TRIzol) and incubated at room temperature for 5 min. Following incubation, 200 μl of chlorophorm was added and the tube was shaken vigorously for 15 s before incubation at room temperature for 2-3 min. The solution was then centrifuged

at maximum speed in a microcentrifuge (cooled at 4°C) for 15 min and the colourless aqueous phase was eluted. Afterwards, 1 μl of (10 mg.ml^{-1}) glycogen and 500 μl of isopropanol were mixed with this eluant, and the reagents were then incubated at room temperature for 10 min. The RNA was then pelleted by centrifugation for 10 min at 4°C, after which the supernatant was removed. The pellet was washed with 1 ml 75% ethanol, air-dried and dissolved in 11 μl DEPC-treated water.

2.9.2 Reverse transcription

cDNA was prepared by reverse transcription using Superscript II reverse transcriptase (GibcoBRL, Life Technologies). Random hexamers were added to the 11 µl containing RNA and heated to 70°C for 10 min. The solution was then cooled on ice and 2 µl 25mM $MgCl_2$, 1 µl 10 mM dNTPs (Pharmacia Biotech) and 2 µl 0.1 M DTT were added to make the reaction, which was mixed and warmed to 42°C for 50 min; the enzyme was then heat inactivated at 70°C for 15 min after which the solution was cooled on ice. Finally, 1 µl of RNAase H (3.8 U.µl^{-1}) was added and the reaction was incubated at 37°C for 20 min to remove the single stranded RNA.

2.9.3 PCR amplification of *VSG* specific cDNA

Three characterised *MVSG* loci, *MVSG1.22*, *1.61* and *1.63*, were examined by PCR for the presence of transcripts both proximal to the promoter and around the *VSG*. PCR was carried out on the cDNAs with ~3 pmoles of upstream and downstream primers with an initial 10 min denaturation at 95°C followed by 30 cycles of 95°C for 1 min, 55°C for 1 min, and 72°C for 1 min. These reactions were carried out using *Taq* DNA polymerase (AB) for each locus except the *MVSG1.61* promoter proximal region, when *Pfu* DNA polymerase was used with 2mM $MgSO_4$ and a lower annealing temperature of 50°C. The PCR primers were designed as follows: *MVSG1.22* promoter proximal, primer 5' ends were +29 and +214 from the transcription start site defined from metacyclic nascent transcripts (Ginger M.L., manuscript in preparation), and *VSG* coding sequence primer 5'

ends were +426 and +666 from the ATG start codon, which is 2934 bp from the transcription start; *MVSG1.61* proximal promoter +28 and +287 from transcriptional start, coding sequence –5 and +305 from ATG, which is 2783 bp from the transcription start; *MVSG1.63* proximal promoter +59 and +230 from transcription start, coding sequence +621 and +837 from ATG, which is 2988 bp from the transcription start.

2.10 Probe manufacture and DNA hybridisation

2.10.1 Radiolabelling

DNA fragments used as probes in this study were originally excised from agarose gels following electrophoresis (using the Qiagen gel extraction kit). The Prime-It II kit was used (Stratagene) for radiolabelling. Initially, 50 ng of purified template DNA were mixed with 10 µl of random oligonucleotides and sterile, distilled water, in a total reaction volume of 37 µl. The mixture was then heated to 95-100°C for 5 min, cooled, and centrifuged briefly. Afterwards, 10 µl 5 x primer buffer, 2 µl α^{32}P labelled dCTP (20 μCi) and 1 µl Klenow (5U.µl^{-1}) were added in order, mixed carefully, and incubated at 37°C for 5 min. The resultant probes were then purified from the unincorporated nucleotides by passing them through Microspin columns (Amersham). Once purified, the probes were denatured at 95 °C for 5 min before use.

2.10.2 Hybridisation

The nylon filters were wetted with distilled water and transferred to a glass hybridisation tube. Approximately 25 ml of Church-Gilbert solution (0.342 M Na_2HPO_4, 0.158 M $NaH_2PO_4.2H_2O$, 0.257 M SDS and 1mM EDTA per litre) was added and the filters were prehybridised for a minimum of 1 h at 65°C in a rotating hybridisation oven. The purified, denatured, probe was then added, and the hybridisation was left overnight at 65°C. After this hybridisation step, the filters were washed at 65°C in the rotating oven with the following series of solutions: 5 x SSC, 0.1% SDS (twice – once using room

temperature solution); 2 x SSC, 0.1% SDS (50 ml solution used per 30 min wash). The filters were then rinsed in 0.1 x SSC (without SDS), heat-sealed in plastic, and placed next to medical photographic film (Konica Medical Corporation) in an autoradiography cassette at –80°C for 4-168 h, depending on the expected strength of the signal. Alternatively the blot was placed next to a phosphoimage screen (Fuji) at room temperature for 1 to 48 h, again depending on the expected strength of the signal.

2.10.3 Stripping of nylon filters

The nylon filters were stripped with boiling 0.1% SDS. The solution was poured onto the filters in a heat resistant container and allowed to cool room temperature. The procedure was then repeated, after which the filter was rinsed in 2 x SSC and was ready for reuse.

2.11 Phenol/chloroform extraction

The volume of the sample was adjusted to 400 µl by the addition of TE buffer. An equal volume of phenol/ chloroform (1:1 mixture) was then added and mixed thoroughly by inversion. The two phases were separated by centrifugation in a microcentrifuge at maximum speed for 10 min, after which the aqueous layer was eluted and transferred to a new eppendorf tube. An equal volume of chloroform was then added and the tube contents were mixed by inversion. After centrifugation at maximum speed for 5 min, the aqueous layer was eluted and added to 1/10 the original volume of 3 M sodium acetate and 2 µl of glycogen (Boehringer Manheim); 2 volumes of 100% ethanol were then added and mixed thoroughly. The tube was then transferred to -20°C for at least 20 min, after which the DNA was pelleted by centrifugation at maximum speed for 15 min. The pellet was then washed in 1 ml 70% ethanol, air-dried, and resuspended in an appropriate volume of buffer (usually TE).

2.12 FISH analysis

FISH analysis was carried out in the laboratory of K. Gull using the following procedure. Slides were coated with silane and attached to Geneframe at RT. The cells were then settled on organosilane-coated slides (5 min). The slides were subsequently washed briefly with PBS and fixed with 4% formaldehyde/5% acetic acid in PBS for 15 min at RT. Two 5 min washes were then carried out with PBS and the cells were permeabilised with 0.1% Nonidet P-40 in PBS for 5 min. The slides were washed in PBS for 5 min. Prehybridisation was carried out for 30 min - 1 h with hybridisation buffer (50 μl/slide, cover with parafilm). Prehybridisation solution was then removed and hybridisation solution including the DNA probe was added at a concentration of about 10 ng/10 μl (using Geneframe which requires a volume of 25 μl). Denaturation was carried out on an *in situ* block for 5 min at 95°C prior to hybridisation overnight at 37°C. Slides were then placed in 50% formamide, 2xSSC for 30 min. The following washes were then carried out: 2xSSC for 10 min at 50°C, 0.2xSSC for 60 min at 50°C and 4xSSC for 10 min at RT. Incubation was performed with the first antibody (anti DIG Fab-fragment, Boehringer, at 0.3 μg/ml in BMEB) for 45 min at 37°C. Three 5 min washes in TBS, 0.05% Tween 20 were carried out prior to incubation with the second antibody (FITC rabbit anti sheep, Vector, at 10 μg/ml in BMEB) for 45 min at 37°C. A further three 5 min washes in TBS, 0.05% Tween 20 were then carried out. Finally the slides were mounted in Vectashield including 100 ng/ml DAPI (99 μl Vectashield plus 1 μl 10 μg/ml DAPI). The hybridisation was composed of 50% formamide, 2xSSC,10% dextran sulphate and 50 mM Na-phosphate buffer; pH 7.

CHAPTER 3

CLONING, CHARACTERISATION AND DISRUPTION OF *KU* IN *T. brucei*

3.1 Introduction

To analyse the function of Ku, initially looking for general phenotypes previously described in other organisms and then questioning possible specific roles in *VSG* switching, we adopted the approach of deleting both alleles of either *KU* to generate null mutants. The strain chosen for this analysis was a bloodstream form MITat 1.2 427 trypanosome strain termed 3174.2 marked with hygromycin and G418 resistance genes that allow analysis of *VSG* switching. However, in order to assess the role of Ku as a mediator of TPE and a potential regulator of differential expression of *MVSG*s, it meant mutating *KU* in a strain containing characterised *MVSG*s. Our laboratory and the laboratory of I. Roditi have tried to transmit *T. brucei* S427 through hundreds of tsetse, to no avail, so were unable to identify *MVSG*s (J.D. Barry, personal communication). Therefore, we instead used *T. brucei* EATRO 795, which routinely develops to the metacyclic stage, permitting characterisation of the *MVSG*s encoding the ILTat 1.22, 1.61 and 1.63 VSGs.

Before trying to analyse the function of Ku in trypanosomes it was necessary to clone both *KU70* and *KU80* from trypanosomes. Due to the availability of limited homologous sequence in *T. brucei* sequence databases, the method chosen to clone either *KU* homologue was the screening of an existing genomic library using PCR-amplified fragments based on database "hits". In order to compare trypanosome *KU* homologues with those from other organisms, it would then be necessary to compile homology alignments with other Ku homologues using Clustal W analysis. This task has been made easier in recent years with definition of the most homologous portions of each protein as primary homology regions (PHRs) (Gell & Jackson, 1999). Gene copy number would also have to be determined before mutation could be carried out.

3.2 PCR amplification of putative *KU* homologue gene fragments

Before full-length copies of each putative *KU* gene homologue could be obtained, it was necessary to amplify and subsequently sub-clone Sanger database DNA sequence homology "hits" for both genes. Cloned products would then be used to probe an ILTat 1.2 genomic library (*Sau*3A partial digests in λ-GEM-12) to allow retrieval of full-length gene sequences along with regions flanking either ORF.

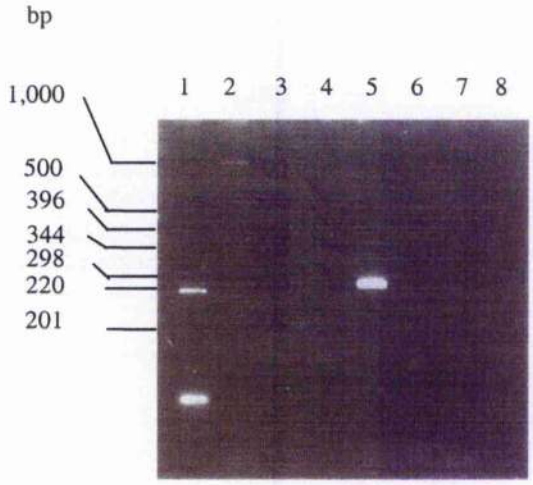

Figure 9. PCR amplification of *KU70* and *KU80* ORF fragments. The fragments were amplified from ILTat 1.2 genomic DNA using oligonucleotides designed from the Sanger clones 0a2 for *KU70* and a27e10 for *KU80* (see Figure 13 below for oligonucleotide positions). Amplification was performed for 30 cycles of 1 min at 95°C, 1 min at 65°C and 1 min at 72°C and a final extension at 72°C for 10 min. The PCR products were size-fractionated on a 0.7% agarose gel. The primers used in each lane were: lane 1, Ku805'P and Ku803'P; lane 2, Ku805'P; lane 3, Ku803'P; lane 4, Ku805'P and Ku803'P (no DNA control); lane 5, Ku705'P and Ku703'P,; lane 6, Ku705'P; lane 7, Ku703'P; lane 8, Ku705'P and Ku703'P (no DNA control). Molecular markers (bp) are located to the left of the panel.

Sanger sequence 0a2, a 248 bp stretch with homology to *KU70*, was amplified using standard PCR conditions (Fig 9). The 5' oligonucleotide used in this reaction was Ku705'P, a 28-mer, 20 bases of which were specific to a region in the 5' region of the

Sanger sequence. This oligonucleotide also contained an *Eco*RI target site, along with 2 bases to allow restriction enzyme binding. The second primer, Ku703'P, was a 29-mer specific to 21 bp at the 3' end of clone 0a2 (see Materials and Methods 2.5.3 for oligonucleotide sequence). A restriction enzyme target site for *Hin*dIII was also located in pKu703'P along with 2 extra bases. After amplification the PCR products were cloned into pBluescript (which had also been digested with *Hin*dIII and *Eco*RI). Selected clones were then screened for insert integration, DNA preparations being digested with *Eco*RI and *Hin*dIII. pKu70(2) and pKu70(3) both liberated a fragment of the requisite size after digestion and were sequenced using the oligonucleotides T3 and T7.

A second Sanger database entry, a27e10, containing a putative 214 bp *KU80* gene homologue fragment, was also amplified by PCR using two specific oligonucleotides (Fig 9). The first oligonucleotide, Ku805'P, was a 26-mer, 20 bases of which were specific to a region at the 5' end of the Sanger sequence. The other 6 bases encoded the recognition sequence for the restriction enzyme *Eco*RI. The second oligonucleotide, Ku803'P, contained a 20 base stretch specific to the 3' end of the Sanger sequence. This oligonucleotide also contained an *Eco*RI restriction enzyme target site (see Materials and Methods 2.5.3 for oligonucleotide sequences). PCR amplification was carried out using standard conditions (see Materials and Methods). The template used for this reaction was 50 ng of ILTat 1.2 genomic DNA. Reaction products were pooled, purified and digested with *Eco*RI. After digestion, samples were purified and cloned into the plasmid vector pBluescript which was also digested with *Eco*RI and purified. Selected transformants were then screened for insertion of the desired product using restriction enzyme mapping. Two clones, pKu80(3) and pKu80(11), were shown to contain an insert of appropriate size and were sequenced using the oligonucleotides T3 and T7.

3.3 Analysis of *KU70* and *KU80* gene-copy number

As gene disruption was the chosen method for analysis of the function of *KU* in *T. brucei,* it was deemed necessary to ascertain whether each protein was encoded by a single or multiple-copy gene. This analysis would also allow detailed mapping of the genomic environment surrounding each gene locus.

ILTat 1.2 genomic DNA was digested with a range of restriction enzymes possessing 6-bp recognition sequences. This was done in duplicate in order to analyse the two genes together. These digests were gel electrophoresed, Southern blotted and probed with pKu70(2) or pKu80(11) specific radio-labelled DNA sequence. Each blot was then washed to 0.2 x SSC, 0.2% SDS, at 65°C. The hybridisation patterns attained following this experiment suggested that both putative *KU70* and *KU80* homologues were probably single-copy genes (Fig 10). This, however, would not be confirmed until deletions of both genes had been attempted.

One interesting observation from this exercise was the hybridisation of either probe to more than one band in some of the DNA digest lanes; *Eco*RV for *KU70* and *Eco*RI and *Eco*RV for *KU80*. Not all these restriction enzymes possess a target site in *KU70* and *KU80*. One explanation for the binding of the probe to two fragments in single digest lanes is that there are allelic differences in the flanks of both genes. However, the fact that both *KU70* and *KU80* are probably single-copy meant that each gene could be disrupted by two independent gene conversion events.

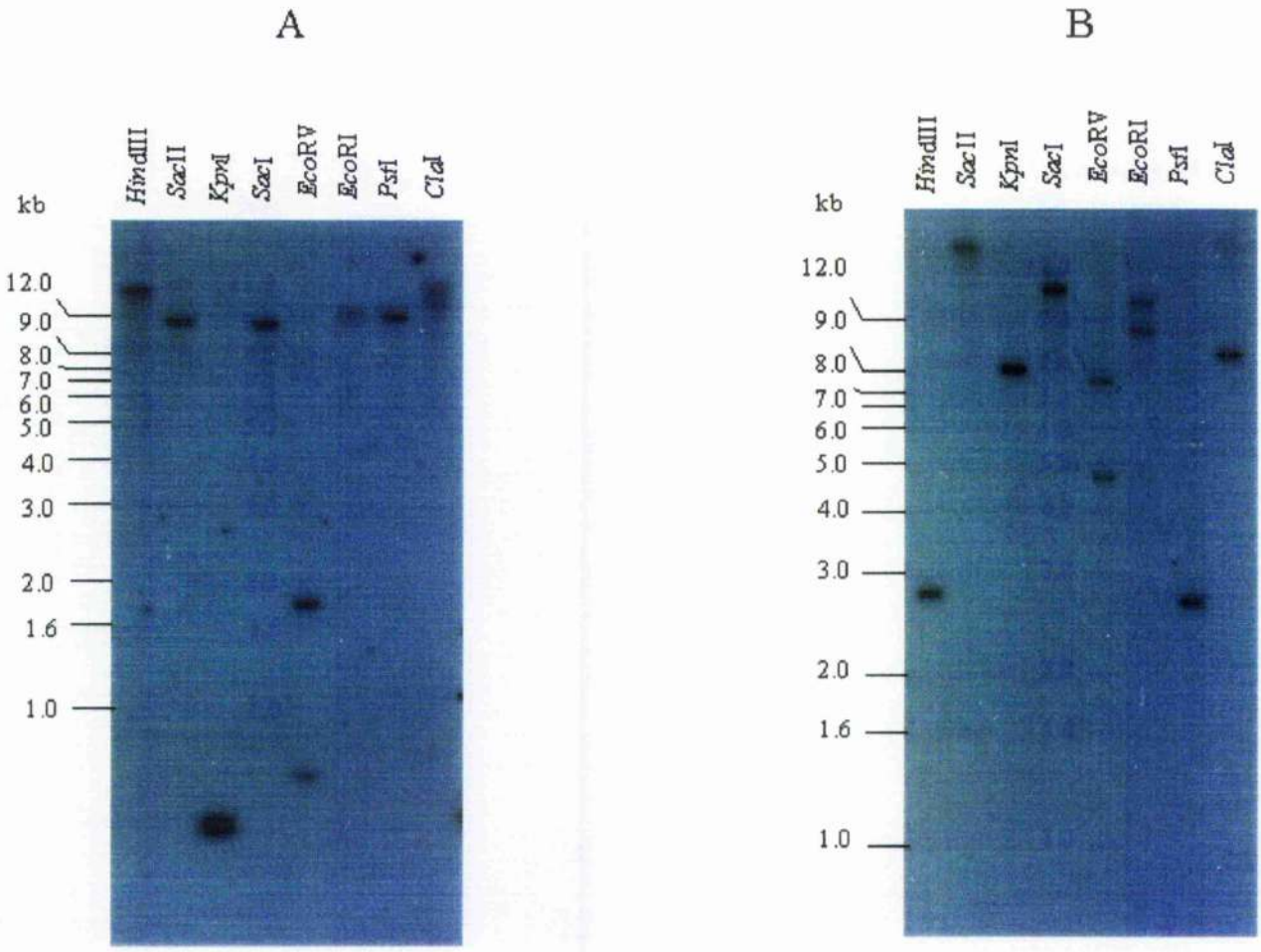

Figure 10. Determination of gene-copy number for putative *KU70* and *KU80*. ILTat 1.2 genomic DNA was digested with a range of restriction enzymes recognizing 6-bp target sites. The digests were size-fractionated on a 0.7% agarose gel and Southern blotted. Filter A was probed with DNA sequence specific to putative *KU70* and filter B was probed with DNA sequence specific to putative *KU80*. Each filter was then washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C. Molecular markers (kb) are indicated to the left of the panel.

3.4 Screening λ-genomic library for putative *KU* homologues

The full-length putative genes were cloned from an ILTat 1.2 genomic DNA library (*Sau3A* partial digests in λ-GEM-12) using DNA sequences obtained from the Sanger database. One advantage of this approach is the avoidance of PCR artefacts. Sub-cloned gene homologue fragments were excised from the pBluescript backbone, using the double-restriction enzyme digest combination of *Pst*I and *Hinc*II for pKU70(2) and *Not*I and *Eco*RV for pKU80(11). The region-specific DNA fragment required to probe the genomic library for the full-length putative homologue of *KU80* was 214 bp, while the partial *KU70* gene fragment was 248 bp. These bands were independently purified using the Qiagen Gel Extraction Kit and 30 ng of each fragment was then α-^{32}P radiolabelled using the Stratagene Prime-It II Kit.

A plate of the genomic library was incubated at 37°C until the plaques were near confluence. The plate was blotted onto Nylon membrane which was then washed with 2 x SSC. A second membrane was blotted using the same plate and was washed in a similar manner. One blot was hybridised at 65°C for 16 hr with α-^{32}P radiolabelled *KU70* gene homologue fragments and washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C. The second blot was hybridised to the radiolabelled *KU80* probe using identical conditions and also washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C. The resultant autoradiographs revealed several positive plaques.

Three λ-clones for each gene were selected for further screening. They were lifted, along with immediately adjacent plaques, using sterile pasteur pipettes, and were eluted in 1 ml of phage buffer. Each of the six eluants was serially diluted and the diluted stocks were plated as in the original screen. After overnight incubation at 37°C, plates containing isolated plaques were blotted and washed as above. Each blot was then hybridised with the same radiolabelled DNA fragment used to probe the original progenitor plate blots. Again, the blots were hybridised at 65°C and washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C. After this secondary screening, autoradiography revealed a significant proportion of positive plaques on each plate, some of which then were extracted and eluted in phage buffer. A final, tertiary screen was

carried out, in the same way. Autoradiographs confirmed that each of the three original λ-clones for each gene carried through three sequential rounds of screening were indeed clonal.

3.5 Genomic analysis of isolated λ-clones

Purified λ-clone DNA preparations (prepared using promega lambdasorb protocol) for both *KU70* and *KU80* were digested with a series of 6-bp cutting restriction enzymes. These digest samples were then size fractionated on a 0.7% agarose gel and transferred to a nylon membrane by Southern blotting. The filter containing bound *KU70* λ-clone digests demonstrated in Figure 11 was probed with a *KU70* specific α-^{32}P radiolabelled probe (same as for library screening, Section 3.4) and washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C. The second filter, containing bound *KU80* λ-clone digest samples, was also probed with α-^{32}P radiolabelled sequence specific DNA and washed to the same final stringency. The radiolabelled sequence utilised as a probe was specific to the 214 bp region of DNA within the *KU80* ORF (same as for *KU80* library screen, Section 3.4).

The *KU70* restriction enzyme digestion pattern demonstrated that the two clones being analysed were physically different (Fig 11A). A lack of suitable restriction enzyme target sites within the sequence used to probe the Southern blot of this gel, along with the need to obtain sequence confirming the candidate gene as a putative homologue of *KU70*, meant that sub-cloning of fragments containing all or part of the gene ORF was a necessity before physical mapping of the whole *KU70* gene could be carried out. The hybridising 3.2 kb *Sac*I fragment of *KU70* λ-clone I was cloned into pBluescript yielding the plasmid pKu70L1. Sequencing with the two DNA specific oligonucleotides, previously used to PCR amplify the fragment used to probe the genomic library, provided approximately 500 bp of sequence, the majority of which overlapped with previously defined sequence. Comparative sequence analysis revealed sufficient sequence homology with *KU70* homologues to warrant further investigation. New oligonucleotides designed using the newly sequenced region yielded further sequence both upstream and

downstream of the library probe DNA sequence (see Materials and Methods 2.5.3). This pattern was repeated until the putative start codon was reached using oligonucleotides upstream of the probe DNA sequence. Oligonucleotides downstream of the probe DNA sequence were used to retrieve sequence towards the C-terminal portion of the putative homologue. However, the cloned DNA insert sequence did not contain the gene terminus: a second λ-fragment (from λ-clone 2), containing the remnant portion of putative *KU70*, would have to be sub-cloned and sequenced.

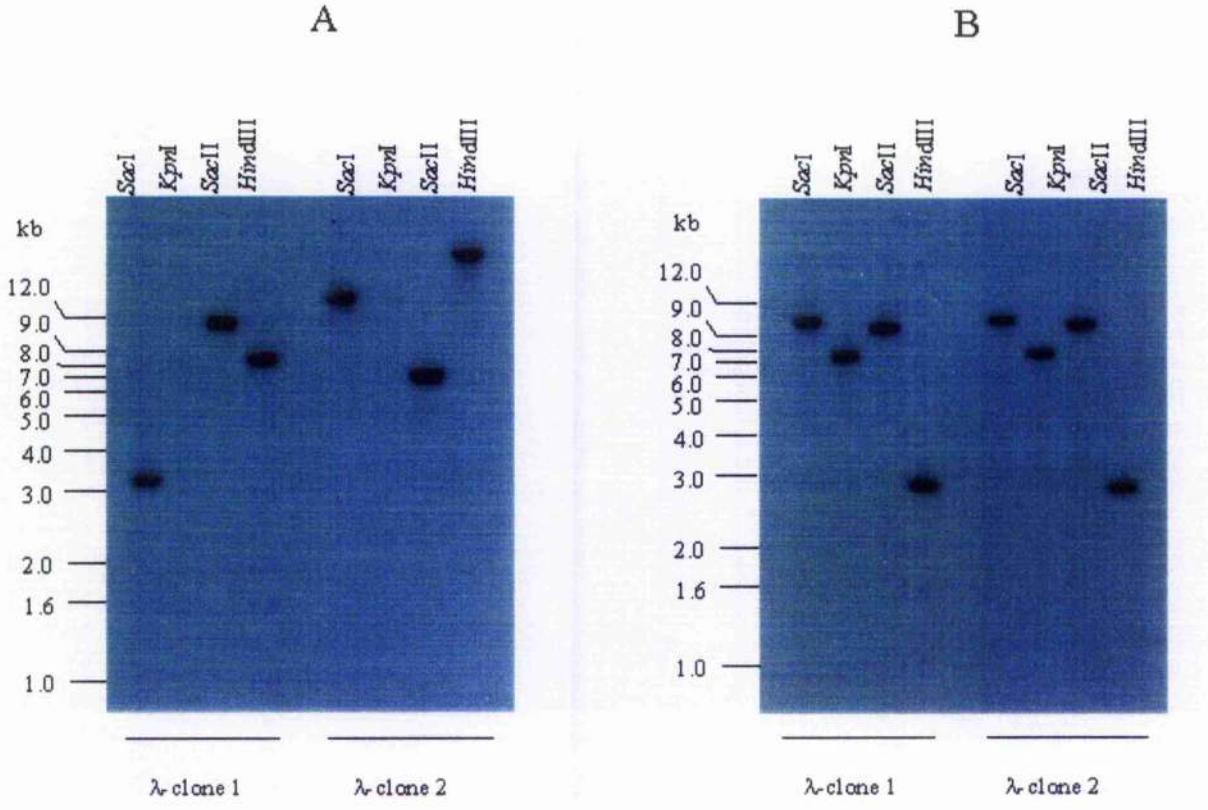

Figure 11. Restriction mapping of *KU70* and *KU80* λ-clones. DNA was extracted from two λ-clones isolated for both *KU70* and *KU80*. The DNA was digested with *Sac*I, *Kpn*I, *Sac*II and *Hin*dIII for the clones containing either *KU70* or *KU80* sequence. The digests were separated on a 0.7% agarose gel and Southern blotted. Filter A is probed with *KU70* specific sequence (excised from pKU70(2)), while filter B is probed with *KU80* specific sequence (excised from pKU80(11)). The lack of *Kpn*I restriction enzyme product hybridising to radiolabelled probe sequence in filter A is most likely due to the small size of the fragment. Both filters were washed to 0.1 x SSC, 0.1% SDS at 65°C. Molecular markers (kb) are indicated to the left of the panel.

A DNA preparation was performed for *KU70* λ-clone 2 (following manufacturer's protocol – Promega). λ DNA was digested with *Hin*dIII and also double digested with *Hin*dIII plus a series of other enzymes that possess a 6-base pair recognition sequence (and do not cut within the previously sequenced portion of *KU70*) (Fig 12). These digests were size-fractionated on a 0.7% agarose gel and transferred to a Nylon membrane by Southern blotting. The filter was then probed to a stringency of 0.1 x SSC, 0.1% SDS at 65°C, with a 248 bp α-^{32}P radiolabelled PCR amplified product previously used to probe the genomic library. *Hin*dIII has a 6 bp recognition sequence approximately 30 bp upstream of the putative *KU70* start and proved useful when trying to orientate physical maps of the putative *KU70* locus.

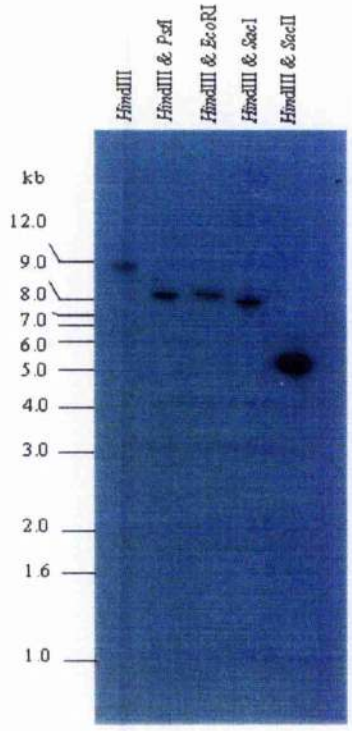

Figure 12. Orientating *KU70* λ-clone to known *Hin*dIII restriction site. DNA was extracted from *KU70* λ-clone (2). The DNA was then digested with *Hin*dIII or *Hin*dIII in conjunction with a range of restriction enzymes. The digests were size-fractionated on a 0.7% agarose gel and the gel was then Southern blotted. The filter was probed with DNA sequence specific to the *KU70* ORF (excised from pKU70(2)) and washed to a final stringency of 0.1 x SSC, 0.1% SDS. Molecular markers (kb) are indicated to the left of the panel.

This hybridisation revealed the presence of various restriction enzyme sites believed to be downstream of the *KU70* ORF. The *Hin*dIII/*Sac*II 5 kb double digestion λ-clone product was sub-cloned into pBluescript. This sub-clone, pKu70L2, was sequenced, allowing us to obtain the previously unidentified putative *KU70* gene stop codon (Fig 13) and gene flank sequence (see Materials and Methods for sequencing oligonucleotides).

The *KU80* λ-clone hybridisations revealed a consistent pattern, confirming that the two clones originally selected were identical (Figure 11B). These hybridisations also revealed fragments of suitable sizes for sub-cloning into a replicative bacterial plasmid. Digestion with *Hin*dIII liberated a fragment of 2.8 kb containing a portion, if not the entirety, of the putative *KU80* ORF. It was deemed appropriate to attempt to sub-clone this particular fragment in order to obtain also further putative *KU80* DNA sequence. Lack of regular 6 bp restriction enzyme recognition sites within the short 250 bp DNA product used to probe these digests left little alternative but to clone the 2.8 kb fragment before orientating the gene ORF within the λ-clone. This would subsequently provide further restriction enzyme target sites to physically map the *KU80* locus.

*Hin*dIII digestion products of *KU80* λ-clone 1 were cloned into pBluescript. DNA preparations were carried out for selected ligation products and the composition of each sub-clone was verified by restriction mapping. Insert DNA for sub-clone pKu80LI was subsequently sequenced using the oligonucleotides used to amplify the DNA fragment that had been used to probe the genomic library. This would reveal not only the partial sequences already obtained, but also some of the intervening trypanosome genomic DNA. Once again, sequence analysis revealed homology between the region sequenced and *KU80* homologue sequence from other eukaryotes. Sequence obtained approximated to 500 bp of DNA per oligonucleotide used. New oligonucleotides were then designed using this sequence, allowing new sequence both upstream and downstream of the library probe DNA sequence to be obtained (see Materials and Methods 2.5.3 for sequencing oligonucleotides). This pattern was continued until the putative start codon was reached. However, the 3' terminus of the gene was absent from the investigated clone. This necessitated further restriction mapping of *KU80* λ-clone 1, followed by sub-cloning of a

restriction fragment containing the remaining *KU80* homologue coding sequence and adjacent flanking region.

A second stretch of sequence encompassing the *KU80* locus was cloned following further mapping, using the *Hin*dIII target site originally used to clone the upstream portion of the gene. Double restriction enzyme digests using *Hin*dIII in conjunction with several other restriction enzymes were carried out for *KU80* λ-clone 1. Digests were size fractionated by gel electrophoresis, Southern blotted onto a Nylon membrane and fixed by UV irradiation. The sequence that was ultimately used to probe this Southern blot was retrieved using PCR amplification. Two specific oligonucleotides were designed, 1O20 forward and 1O20 Reverse (obtained from the TIGR database entry 1O20; see Figure 13), so that they would prime amplification from loci equidistant from this *Hin*dIII recognition site. The PCR reaction used standard conditions. The reaction products were analysed by gel electrophoresis and the dominant 120 bp product was isolated by gel extraction using the Qiagen Gel Extraction Kit. An α-^{32}P radiolabelled probe of this sequence hybridised, at a final stringency of 0.1 x SSC, 0.1% SDS at 65°C, to the two fragments, one either side of the *Hin*dIII target site produced by the double digests (see Fig 14), giving a comprehensive restriction enzyme map around this specific *Hin*dIII target site. It was decided that the 2.4 kb *Hin*dIII/*Sac*I double digest product would be cloned and then sequenced, thus allowing retrieval of remnant putative *KU80* gene sequence along with flanking sequence adjacent to its gene stop codon.

KU80 λ-clone 1 was digested with *Hin*dIII and *Sac*I. pBluescript was digested with the same restriction enzyme combination using standard reaction conditions and both plasmid and insert were purified prior to cloning. The resulting sub-clone pKu80L2 was analysed by restriction enzyme mapping and subsequently sequenced (Fig 13).

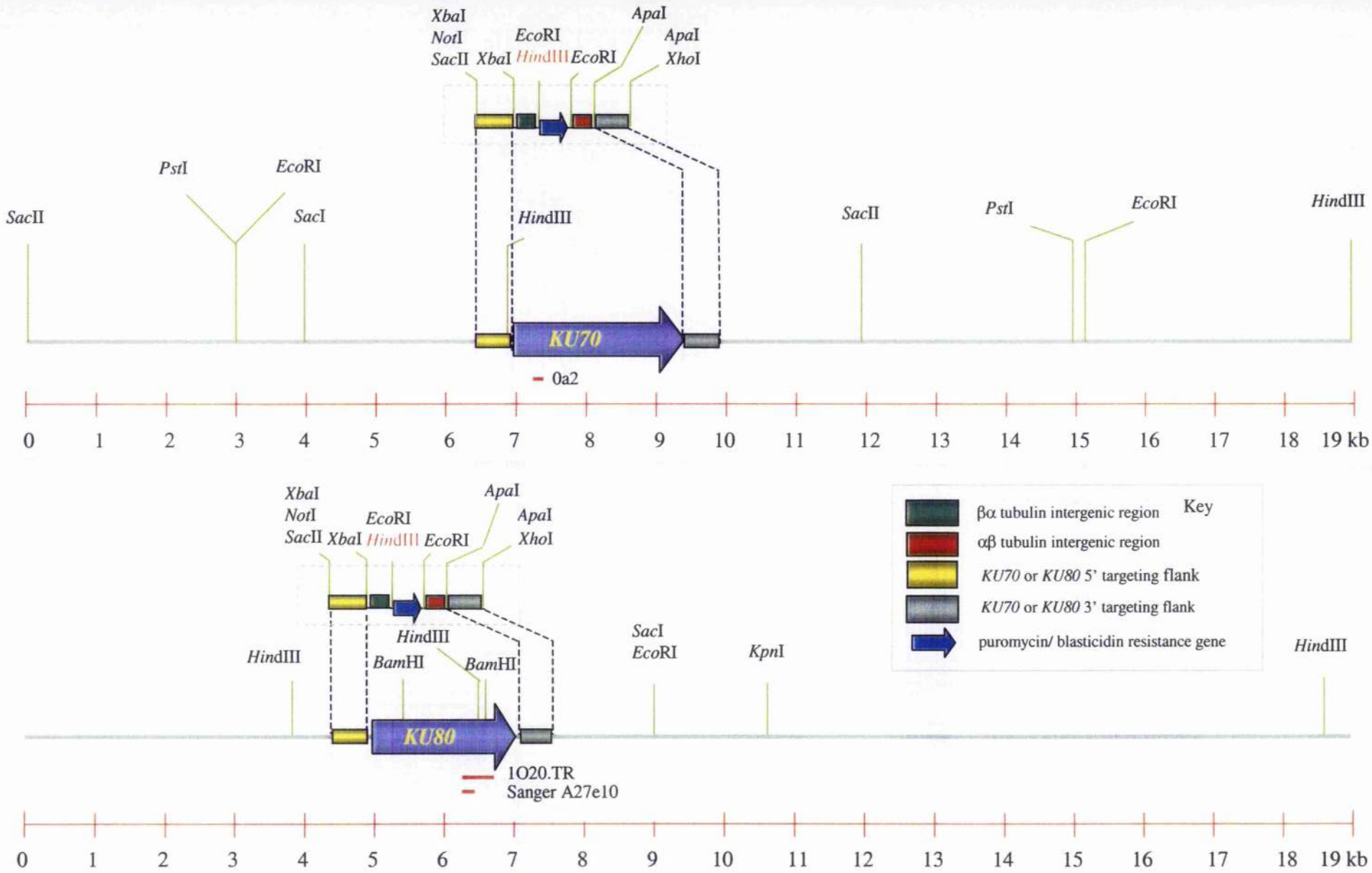

Figure 13. Restriction maps of wild-type and recombinant *KU* loci. Each *KU* gene was sequenced on both strands. These maps were prepared following Southern blot analysis. Genome sequence database "hits" originally used to retrieve full length copies of each gene are highlighted using red lines. Cassettes used to disrupt *KU* loci are represented within dash style grey boxes (see Fig 16). *KU* open reading frames are represented using purple filled arrows. *HindIII* = Restriction site present in pTPT but not in pTBT (see Fig 16).

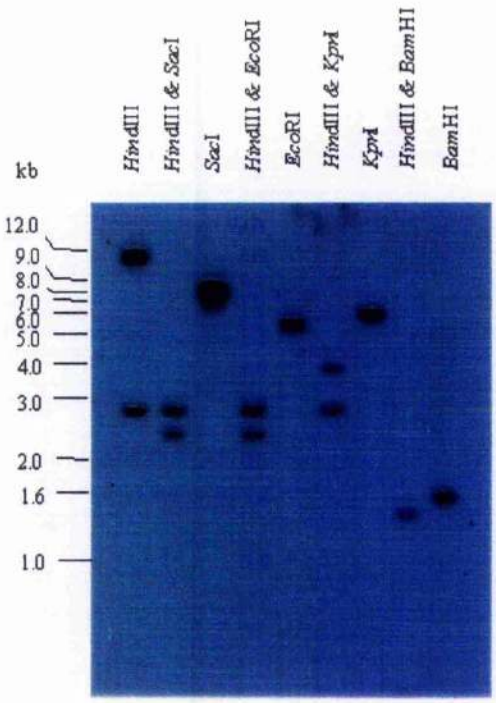

Figure 14. Detailed restriction mapping of *KU80* λ-clone 1. DNA was extracted from *KU80* λ-clone 1, digested with *Hin*dIII in conjunction with a number of other 6-bp cutters. The digests were separated on a 0.7% agarose gel and Southern blotted. The blot was probed with radiolabelled sequence encompassing 60 bp either side of the previously mapped *Hin*dIII restriction enzyme site. Following hybridisation the filter was washed to 0.1 x SSC, 0.1% SDS, at 65°C. Molecular markers (kb) are indicated to the left of the panel.

3.6 Trypanosome Ku homologues are not colinear

As neither Ku70 nor Ku80 has a high degree of interspecies similarity even at the polypeptide level (42), the *T. brucei* genome sequence databases were blast-searched individually with all known homologues. Although most hits were unique to each query sequence, some were shared by more than one. Detailed searching of databases, using Smith-Waterman algorithms, supported the view that these common hits did indeed correspond to Ku homologues. Sequence analysis revealed complete open reading frames (trypanosomes are almost fully intronless) potentially encoding polypeptides of 81 kDa for the Ku70 homologue and 69 kDa for the Ku80 homologue. It was possible to align

parts of those sequences with the primary homology regions (PHRs) of other Ku homologues by Clustal W analysis (Fig 15). The predicted sizes of the trypanosome Ku proteins are the inverse of what occurs in other organisms, and their designation is complicated further by the general relatedness of Ku70 and Ku80 and the dispersion of homologous sequences throughout the proteins. Our belief that the larger protein corresponds to Ku70 relies on the outcome of blastp and PSI-blast searching, which routinely place it closer to Ku70 than Ku80 proteins, and the presence at its carboxy terminal of a partial putative DNA binding SAP domain (6;43) that has the secondary structure potential of other SAP sequences. SAP is a small DNA binding domain with its core strongly predicted to contain two helices separated by a relatively long extended region (Aravind and Koonin, 2000). Trypanosome Ku80 likewise detects mainly Ku80 sequences in blast searching. It does not have a recognisable DNA-PK binding sequence found at the end of vertebrate Ku80 proteins, but nor do the yeast Ku80 proteins.

3.7 Generation of constructs to disrupt *KU* ORFs

To study the function of the Ku heterodimer (composed of Ku70 and Ku80) in *T. brucei*, it was deemed necessary to disrupt each putative gene homologue. It was decided that the initial investigations would be carried out in bloodstream form trypanosomes, with subsequent analysis in procyclic form cells. This would allow us to test whether Ku is involved in *VSG* switching.

Due to the lack of knowledge of the location of important domains in each Ku protein, it was decided that the whole ORF of each gene would be disrupted by targeted gene conversion. To achieve this, two separate cassettes were generated for each gene; one to disrupt each allele. Each cassette contained targeting flanks utilising DNA sequences immediately outside the gene ORF. Antibiotic resistance genes were used as selectable markers, with adjacent tubulin processing signals allowing gene expression after targeted disruption (Fig 16).

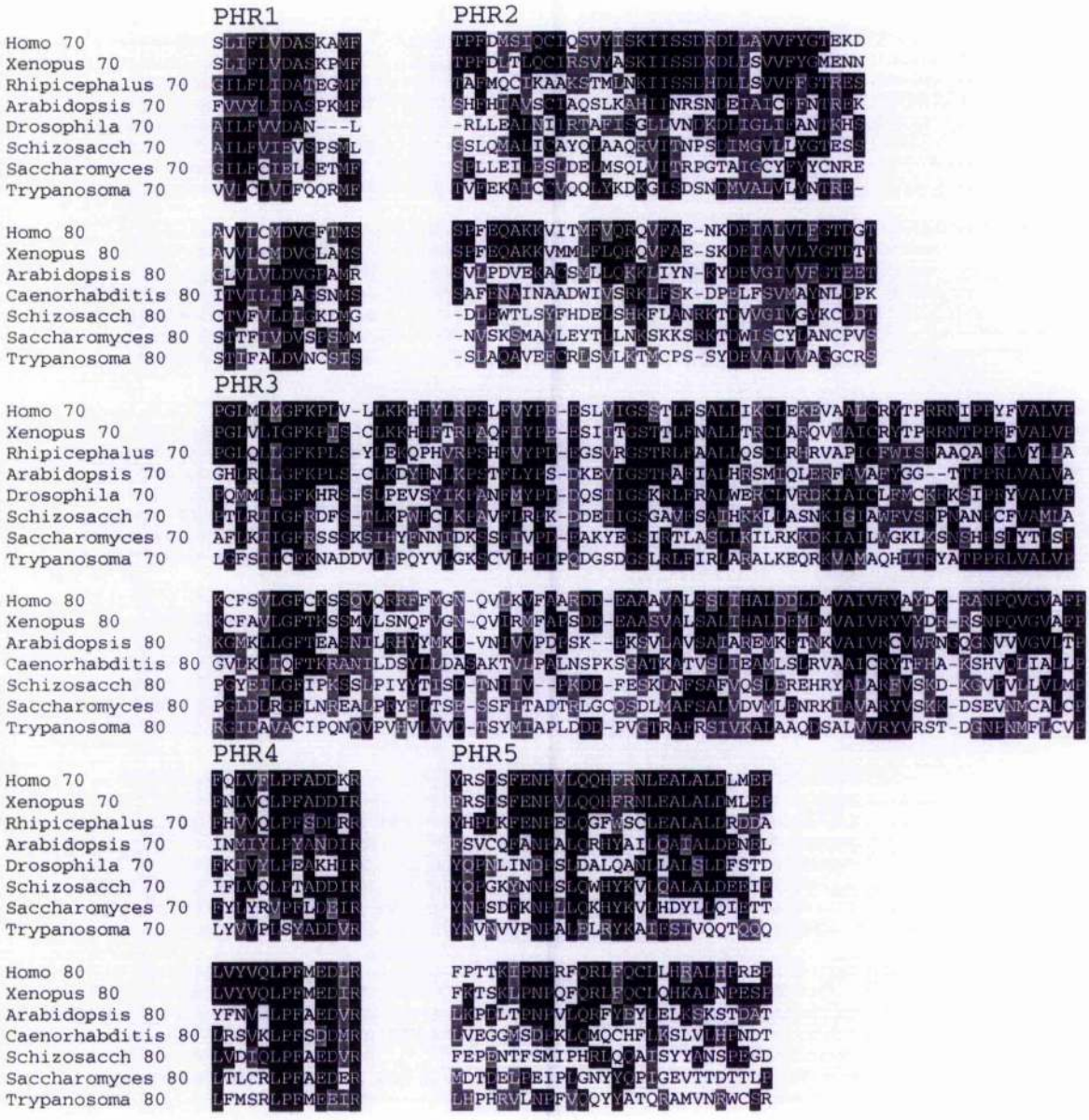

Figure 15. Alignment of the amino acid sequences of eukaryotic Ku proteins. Amino acid residues are shown in single letter code, and the predicted *T. brucei* polypeptides are compared with Ku70 and Ku80 polypeptide sequences from *Homo sapiens*, *Xenopus laevis*, *Rhipicephalus appendiculatus*, *Arabidopsis thaliana*, *Drosophila melanogaster*, *Schizosaccharomyce pombe* and *Saccharomyces cerevisiae*. Residues that are identical for either set of polypeptides are shown in a black background, with conserved residues represented in a grey background. These alignments depict only the predicted Primary Homology Regions (PHRs); PHR1 to 5 are residues 49-61, 94-128, 444-518, 539-551 and 584-611 respectively for *T. brucei* Ku70, and 6-18, 18-53, 327-399, 408-420 and 462-489 for *T. brucei* Ku80.

For both knockout constructs, pTBT70KO and pTPT70KO, the 5' targeting flank was PCR amplified using DNA specific primers (see Materials and Methods 2.6.1). The resultant DNA molecules were digested with *Xba*I and purified. Plasmids pTBT and pTPT, containing genes conferring resistance to blasticidin and puromycin respectively, were digested with *Xba*I. The products were purified and ligated to the targeting flank fragment. DNA preparations from transformants were screened for integration of the 5' targeting flank. Three clones found to contain the targeting flank were sequenced for confirmation of correct integration. Clones containing the correct insert were called pTBT705'F or pTPT705'F, depending on their drug resistance marker genes.

The 3' targeting flank was amplified by PCR. The first primer 703F5P was a 30-mer, 20 bases of which were specific to a region adjacent to the *KU70* ORF (oligonucleotides in section 2.6.1). This region lies immediately downstream of the gene stop codon. The other 10 bases contained the recognition sequence for the restriction enzyme *Apa*I along with an extra 4 bases. The downstream primer 703F3P was a 40-mer containing 20 bases of identical sequence to a region approximately 500 bp downstream of the gene stop codon. This oligonucleotide also contained the recognition sequence for *Apa*I and *Xho*I. The downstream targeting flank was PCR amplified using standard amplification conditions. ILTat 1.2 genomic DNA was used as template for this reaction. The desired product was approximately 500 bp long and, following electrophoretic separation, the product was purified and digested with *Apa*I. Plasmids pTBT705'F(1) and pTPT705'F(1) were also digested with *Apa*I and the 3' targeting flank was subsequently sub-cloned into each plasmid. Sequence analysis of the chosen clones revealed the orientation of the targeting region in the construct. Clones used to disrupt *KU70* were called pTBT70KO and pTPT70KO (Fig 16).

The strategy used to generate constructs capable of disrupting *KU80* was identical to that described in the previous section. The 5' targeting flank lies immediately upstream of the putative *KU80* gene ORF and was amplified using 2 specific oligonucleotide primers to amplify sequence approximately 500 bp upstream of the *KU80* start codon (see Materials and Methods 2.6.1). The two primers were used to amplify the requisite targeting flank using standard amplification conditions. The products were then sub-cloned into the *Xba*I

site of both pTBT and pTPT. Plasmid DNA was prepared from clones selected by antibiotic resistance and three clones containing the correct insert in the required orientation were attained following restriction mapping. Once more, the region encompassing the PCR amplified flank was sequenced for the mapped clones and compared with the template sequence from which it was originally amplified. Clones were called pTBT805'F and pTPT805'F depending on the drug resistant marker within the targeting cassette.

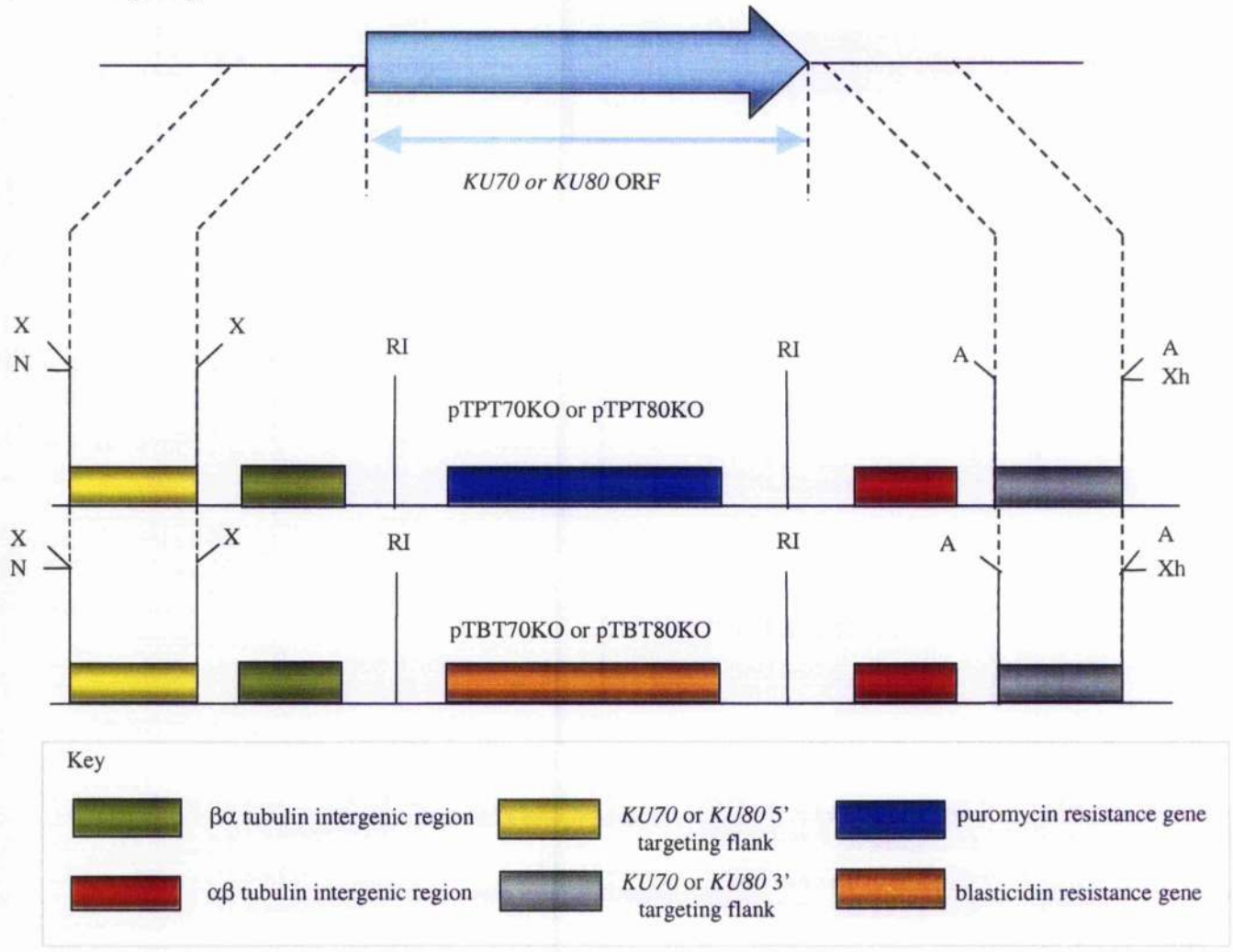

Figure 16. Targeting cassettes generated to disrupt either *KU* ORF. Targeting flanks were were amplified by PCR from ILTat 1.2 genomic DNA and lie immediately outside either ORF. The whole ORF for *KU70* or *KU80* is replaced by expression cassettes that give blasticidin (pTBT70KO or pTBT80KO) or puromycin (pTPTKO70 or pTPTKO80) resistance; these contain 5' and 3' sequences from the tubulin locus that allow transcripts of the resistance genes to be processed into mature mRNAs. Restriction enzyme recognition sites, (N) *Not*I; (X) *Xba*I; *Eco*RI; (A) *Apa*I; (Xh) *Xho*I.

The 3' flank was amplified in a similar manner, using two specific oligonucleotides to amplify approximately 400 bp of sequence immediately downstream of the ORF stop codon. The reaction to amplify the targeting flank used standard conditions. The product was sub-cloned into both plasmids pTBT805'F and pTPT805'F. The region encompassing the amplified product was then sequenced and finally compared with template DNA sequence. The plasmids ultimately used to disrupt both *KU80* alleles were termed pTBT80KO and pTPT80KO (see Fig 16).

3.8 Generation of heterozygous mutants in S427 bloodstream form trypanosomes

Electroporation of linear double stranded DNA was used to disrupt both alleles for each gene in a marked BF MITat 1.2 427 trypanosome strain termed 3174.2 by targeted gene conversion. Plasmids pTBT80KO and pTPT70KO were digested with *Not*I and *Xho*I, releasing the targeting cassette.

A total of $5x10^7$ trypanosomes were transformed per electroporation, 5 μg of purified plasmid digest being used for each transformation. Generation of two independent heterozygous mutant clones for both *KU70* and *KU80* was achieved utilising several transformations for each plasmid. Electroporated cells were then allowed to recover for 16 h at 37°C in 10 ml HMI-9 medium. Viable transformants were selected on semi-solid HMI-9 agarose plates, 10^7 cells being seeded onto any one plate. Plates were subsequently incubated at 37°C for 96-144 hrs. Trypanosomes electroporated with pTBT80KO were selected for stable integration of the targeting cassette using 10 μg.ml^{-1} blasticidin per plate, while cells transformed with pTPT70KO were selected on HMI-9 plates containing 1 μg.ml^{-1} puromycin. Following growth on selective medium, trypanosome colonies were eluted into 10 ml HMI-9 liquid medium containing the appropriate antibiotic, allowing confirmation of drug resistance.

3.9 Genomic analysis of 1st round transformants (bloodstream form)

Genomic DNA was prepared for each original independent transformant and preparations were digested with *Sac*II for cells transformed with pTPT70KO and *Hin*dIII for those transformed with pTBT80KO. The digests were fractionated on a 0.7% agarose gel, Southern blotted onto a nylon membrane and bound by UV irradiation. Fig 17A represents transformants derived from 3174.2 wild-type cells electroporated with the targeting cassette released from pTPT70KO by restriction enzyme digestion. This blot was probed with the α-^{32}P radiolabelled downstream targeting flank of pTPT70KO and washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C.

The radiolabelled targeting flank used to probe the Southern blot in Figure 17A hybridised to a single fragment in the *Sac*II digest of 3174.2 wild-type genomic DNA. Correct integration of the cassette excised from plasmid pTPT70KO should result in hybridisation of the probe to both the remnant wild-type restriction fragment and a 3 kb fragment resulting from the disrupted allele. Four out of 5 transformants analysed exhibit this hybridisation pattern and thus appear to represent true heterozygous mutants. Clones 70P3.1+/- and 70P4.2+/- (generated from two separate transformations) were used for further experiments.

A Southern blot of genomic DNA digests from 3174.2 wild-type cells transformed with the targeting cassette from pTBT80KO is shown in Figure 17B. This blot was probed with the upstream targeting flank of the plasmid pTBT80KO and after hybridisation the blot was also washed to a final stringency of 0.1xSSC, 0.1% SDS at 65°C. The radiolabelled 5' targeting flank, used to probe the Southern blot in Figure 17B, hybridised to a single fragment in the *Hin*dIII digest of 3174.2 wild-type genomic DNA. Correct integration of the first *KU80* allelic disruption cassette should result in hybridisation of the probe to an 11 kb fragment in conjunction with hybridisation to the 3 kb wild-type restriction fragment. This pattern is seen in all lanes, for both 80B1.1+/- and 80B1.3+/-.

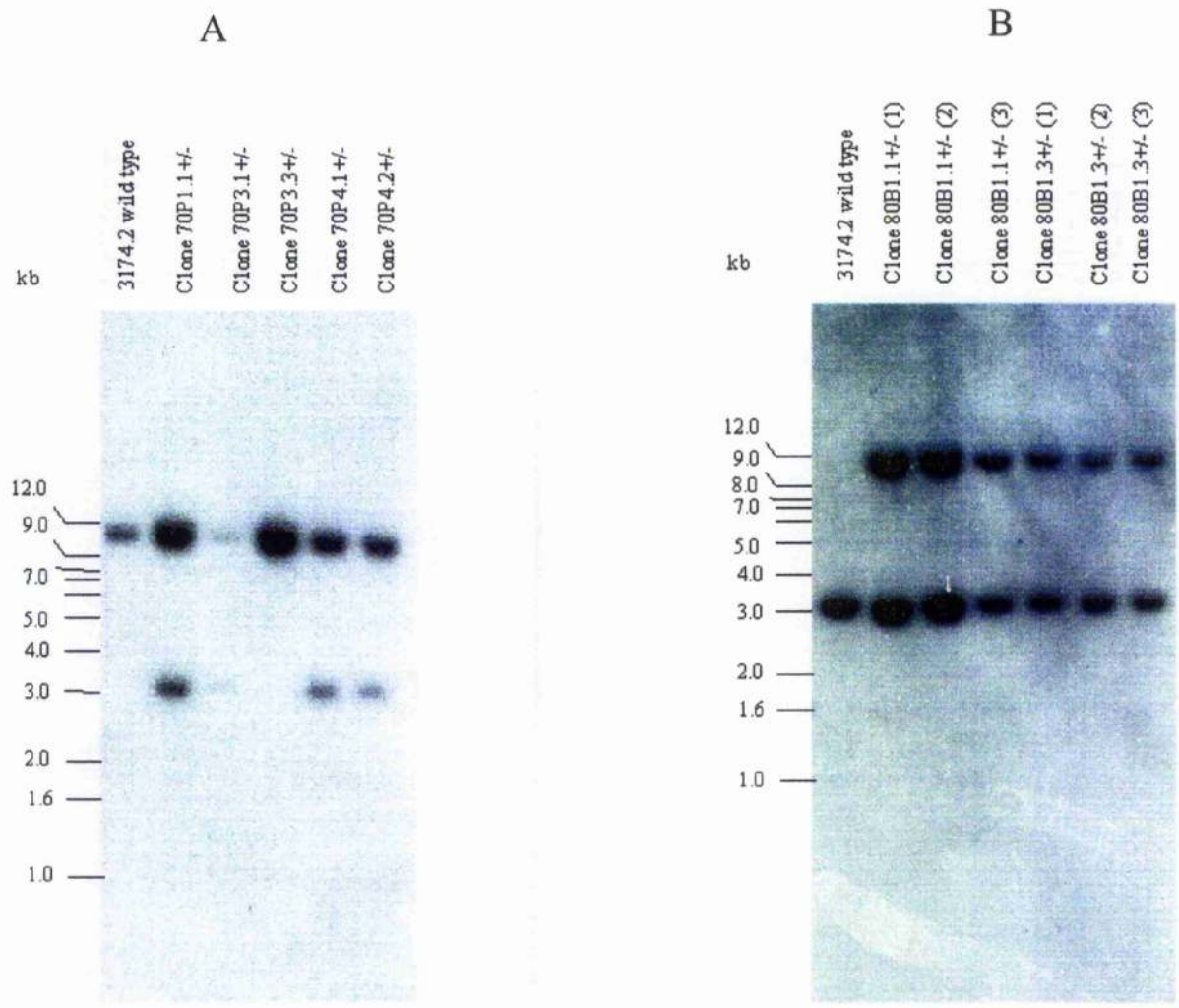

Figure 17. Characterisation of heterozygous bloodstream form mutants. Genomic DNA was extracted from wild-type 3174.2 trypanosomes and first round transformant clones 70P1.1+/-, 70P3.1+/-, 70P3.3+/-, 70P4.1+/- and 70P4.2P+/- (70P3.1+/- and 70P3.3+/- represent clones from the same transformation as do 70P4.1+/- and 70P4.2P+/-). DNA was digested with *Sac*II, size-fractionated on a 0.7% agarose gel and Southern blotted. The blot was then probed with radiolabelled DNA sequence immediately downstream of the *KU70* ORF and washed to a final stringency of 0.1 x SSC, 0.1% SDS, at 65°C. (B) Genomic DNA was extracted from wild-type 3174.2 trypanosomes and first round transformant clones 80B1.1+/-(1-3) and 80B1.3+/-(1-3) (numbers in brackets represent clone numbers isolated after a single transformation; 80B1.1+/- and 80B1.3+/- are two independent transformant clone groups). DNA was digested with *Hin*dIII, size-fractionated on a 0.7% agarose gel and Southern blotted. The blot was then probed with radiolabelled DNA sequence immediately upstream of the *KU80* ORF and washed to a final stringency of 0.1 x SSC, 0.1% SDS, at 65°C. Molecular markers (kb) are indicated to the left of the autoradiograph.

3.10 Generation of *ku* homozygous mutants

Clones representing two independent *KU70* and *KU80* heterozygous mutants were grown in HMI-9 medium, until there were 5 x 10^7 cells in total per trypanosome line. Clones 70P4.2+/- and 70P3.1+/- (Fig 17A) were then transformed with *NotI/XhoI* digested pTBT70KO, while clones 80B1.1+/- and 80B1.3+/- (see Fig 17B) were transformed with 5 μg of purified *NotI/XhoI* digested pTPT80KO. Two separate electroporations were carried out for both clones and the cells allowed to recover for 16 h in 10 ml HMI-9 medium. Transformants were then selected on semi-solid HMI-9 plates as in Section 3.8, this time selecting with 0.5 μg.ml^{-1} puromycin and 5 μg.ml^{-1} blasticidin. Once again, trypanosome colonies were given 96-144 h to grow through selection. Colonies were then eluted in 10 ml HMI-9 medium containing full concentrations of both antibiotics (puromycin at 1.0 μg.ml^{-1} and blasticidin at 10 μg.ml^{-1}) confirming resistance. Genomic DNA was prepared for each independent clone prior to genomic Southern analysis.

3.11 Genomic analysis of second round transformants

Genomic DNA from *KU70* second round transformants was digested with *Sac*II and gel electrophoresed along with wild-type and heterozygous genomic DNA digests. Once again, the gel was Southern blotted onto a Nylon membrane and the DNA was bound by UV radiation. This blot was probed with *KU70* 3' flanking sequence which was radiolabelled and then washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C. This blot demonstrated the loss of expected wild-type allelic restriction fragments. Confirmation of this result is seen in Figure 18 when genomic DNA from *KU70* wild-type, *KU70* heterozygous and *ku70* homozygous cells was digested with *Eco*RI, Southern blotted and probed with radiolabelled *KU70* ORF sequence. There is no hybridisation to the wild-type band in the homozygous mutant genomic DNA lanes (Fig 18). The integrity of the DNA used in this experiment was tested after this blot was stripped and probed with radiolabelled *VSG221* sequence (Chapter 4, Figure 22).

Genomic DNA prepared from *KU80* second round transformant clones, along with 3174.2 wild-type gDNA and gDNA from the heterozygous mutants, was digested with *Hin*dIII, gel electrophoresed and Southern blotted. Probing with ^{32}P radiolabelled *KU80* 5' flanking sequence, to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C, revealed the absence of hybridisation to the 3 kb wild-type restriction fragment by all 2nd round transformant lanes except clones 80B1.1P -/- (2) and 80B1.1P-/-(7) (Fig 19). Clones 80B1.1P -/- (2) and 80B1.1P-/- (7) were not used in subsequent experiments.

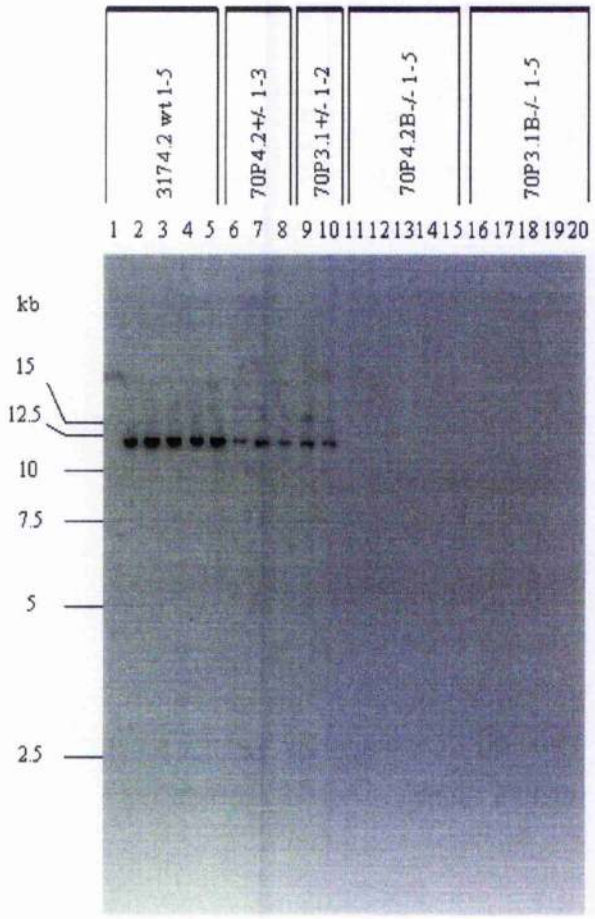

Figure 18. Characterisation of homozygous *ku70* bloodstream form mutants. Genomic DNA was extracted from 3174.2 wild-type, heterozygous and second round transformant clones (used for TLM assay – Chapter 4). DNA was digested with *Eco*RI, size-fractionated on a 0.7% agarose gel and Southern blotted. The blot was then probed with radiolabelled DNA sequence in the *KU70* ORF and washed to a final stringency of 0.1 x SSC, 0.1% SDS, at 65°C. Molecular markers (kb) are indicated to the left of the membrane.

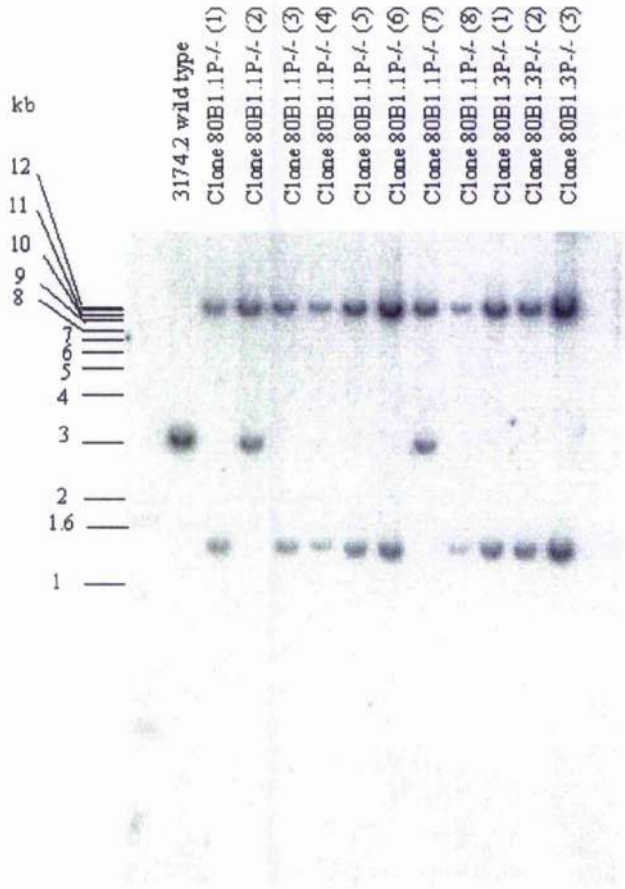

Figure 19. Characterisation of homozygous *ku80* bloodstream form mutants. Genomic DNA was extracted from wild type 3174.2 trypanosomes and second round transformant clones. DNA was digested with *Hin*dIII, size-fractionated on a 0.7% agarose gel and Southern blotted. The blot was then probed with radiolabelled DNA sequence immediately upstream of the *KU80* ORF (excised from pTBT80KO) and washed to a final stringency of 0.1 x SSC, 0.1% SDS, at 65°C. Molecular markers (kb) are indicated to the left of the membrane.

3.12 Generation of heterozygous procyclic form *ku* mutants

Once again, electroporation of double-stranded DNA was used to disrupt both alleles encoding *KU70* in procyclic form EATRO 795 trypanosomes. A total of 3 x 10^7 trypanosomes were transformed with 5 µg of purified pTPT70KO digest per electroporation. Generation of two independent heterozygous mutant clones was achieved by two separate electroporations per plasmid type. Electroporated cells were then allowed to recover for 24 h at 37°C in 10 ml of SDM medium supplemented with heat inactivated foetal calf serum. Serial dilutions of recovered cells were then selected on 1 $\mu g.ml^{-1}$ puromycin and selected populations were then cloned in 96-well plates using conditioned medium.

3.13 Genomic analysis of 1st round transformants

Genomic DNA was harvested from transformants selected on puromycin and each DNA sample was digested with *Sac*II and fractionated on a 0.7% agarose gel along with an EATRO 795 wild-type *Sac*II genomic DNA digest sample. The gel was Southern blotted onto a Nylon membrane and the DNA was bound by UV irradiation. The blot was then probed with ^{32}P radiolabelled 3' targeting flank (excised from pTBT70KO) and washed to a final stringency of 0.1 x SCC, 0.1% SDS at 65°C.

The targeting flank used to probe the blot hybridised to a single *Sac*II fragment in the wild-type genomic DNA digest lane (Fig 20). Correct integration of the cassette used to disrupt a single *KU70* allele should result in hybridisation of the probe to a 3 kb fragment in conjunction with hybridisation to the wild-type fragment. This pattern is seen in lanes 2-4 of Figure 19. This indeed suggested that all transformants analysed had a successfully disrupted *KU70* allele.

3.14 Generation of procyclic form *ku70* homozygous mutants

KU70 heterozygous mutants 70 1.1+/- and 70 1.2+/- (see Fig 20) were independently transformed with 5 μg of *Xho*I/*Not*I digested pTBT70KO. Once again, 3 x 10^7 trypanosomes were electroporated per transformation, utilising two transformations per cell line. Electroporated cells were allowed to recover for 24 h and were selected as in section 3.12 this time with both puromycin (0.5 $\mu g.ml^{-1}$) and blasticidin (5 $\mu g.ml^{-1}$). Populations growing through selection were cloned using the same strategy as used for the heterozygous transformants.

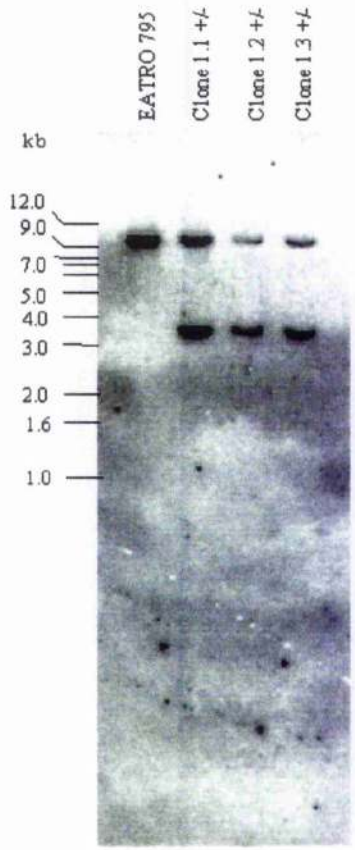

Figure 20. Characterisation of heterozygous *KU70* procyclic form mutants. Genomic DNA was extracted from transformant clones 1.1-1.3 and *T. brucei* EATRO 795. DNA was digested with *Sac*II, size-fractionated on a 0.7% agarose gel and Southern blotted. The blot was then probed with radiolabelled DNA sequence immediately downstream of the *KU70* ORF (excised from pTBT70KO) and washed to a final stringency of 0.1 x SSC, 0.1% SDS, at 65°C. Molecular markers (kb) are indicated to the left of the membrane.

3.15 Genomic analysis of second round transformants

Genomic DNA was prepared from transformants 1.1-/-(5), 1.1-/-(6), 1.2-/-(1) and 1.2-/-(2). Transformant genomic DNA samples, heterozygous clone samples and wild-type genomic DNA were independently digested with *Sac*II (1 μg per sample type). Digest samples were then size-fractionated on a 0.7% agarose gel. The gel was Southern blotted and probed with ^{32}P radiolabelled *KU70* gene ORF DNA sequence and washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C. Figure 21A demonstrates that only heterozygous and wild-type DNA samples hybridised to the ORF probe, the transformants under analysis appearing to have lost both wild-type alleles. This blot was subsequently stripped and probed with radiolabelled *RAD51* ORF sequence. This confirmed that none of the genomic DNA samples bound to the blot were degraded (Figure 21B).

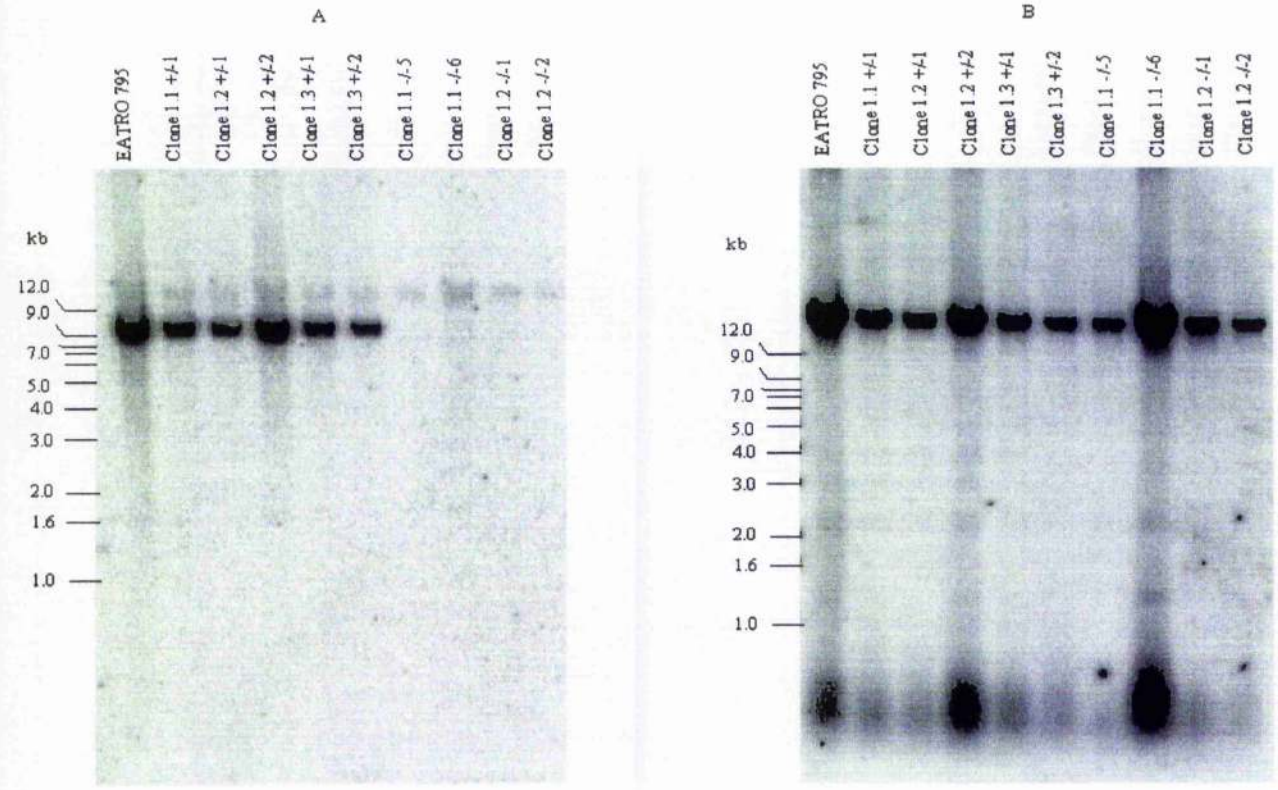

Figure 21. Characterisation of homozygous *ku70* procyclic form mutants. (A) Genomic DNA was extracted from heterozygous clones 1.1-1.3, EATRO 795 and second round transformant clones 1.1-/-5, 1.1-/-6, 1.2-/-1 and 1.2-/-2. DNA was digested with *Sac*II, size-fractionated on a 0.7% agarose gel and Southern blotted. The blot was then probed with radiolabelled DNA sequence from within the ORF of *KU70* and washed to a final stringency of 0.1 x SSC, 0.1% SDS, at 65°C. (B) The latter blot was subsequently stripped and probed with radiolabelled *RAD51* ORF sequence. This was washed to the same stringency. Molecular markers (kb) are indicated to the left of the filters.

CHAPTER 4

PHENOTYPIC ANALYSIS OF *ku* MUTANTS AND THE ROLE OF *KU* IN DIFFERENTIAL *VSG* EXPRESSION

4.1 Introduction

Analysis of the function of *KU* in trypanosomes meant developing a variety of assays to monitor the effect of mutating the homologue for either *KU70* or *KU80*. Assays to analyse the effect of disrupting both genes were developed utilising techniques pioneered in other eukaryotic systems.

Initial experiments were to determine whether the putative *KU* orthologues were authentic. Upon confirmation of orthologue status, *T. brucei* mutants for both genes were tested for their ability to switch the *VSG* gene being expressed in the mammalian host. The hypothesis underlying this work is that Ku is central in governing differential *VSG* expression. We propose that Ku is involved directly or indirectly in blocking gene duplication reactions that cause *VSG* switching in the mammalian host. This role is likely to be extremely important in pleomorphic trypanosome strains with high frequencies of recombination-based *VSG* switching reactions. The work highlighted here however, has been carried out in a laboratory adapted monomorphic strain manipulated to function as a model for *VSG* switching reactions. Ku is also involved in causing a phenomenon in budding yeast termed the telomere position effect (TPE). Genes present upstream of telomere repetitive tracts are subject to transcriptional silencing. Upon disruption of Ku, these genes become transcriptionally active in *S. cerevisiae* (Boulton & Jackson, 1998) but remain inactive in *S. pombe* (Manolis *et al.*, 2001). While TPE has been demonstrated mainly in laboratory conditions, where reporter genes have been inserted artificially in telomeres (Gottschling *et al.*, 1990; Baur *et al.*, 2001), there is little evidence that it occurs naturally. It is therefore possible to invoke Ku, as a candidate for regulating differential expression and repression of the distinct set of *VSG*s employed when the VSG coat first appears, in the metacyclic trypanosome stage in the salivary glands of the tsetse (reviewed in Barry *et al.*, 1998; Donelson *et al.*, 1998).

To determine whether disruption of *KU70* and *KU80* had a fundamental effect on the cells under investigation, growth rates of mutants were followed and compared with wild-type trypanosomes. This was carried out in conjunction with growth assays in the presence of DNA damaging agents to determine cell sensitivity to the occurrence of DNA

breaks. Two separate approaches to this were used. The first looked at populations, measuring the growth rate of cell lines under analysis at various time-points with or without the presence of the partially radiomimetic alkylating agent, methyl methanesulphonate (MMS). A more stringent assay, also utilising MMS, measured the ability of trypanosome sub-populations to grow in the presence of the DNA damaging agent, relative viability being measured by the percentage of wells containing dividing populations after approximately 7 days.

The Ku heterodimer is found bound to telomeres in *Saccharomyces cerevisiae* prior to induction of DNA double-stranded breaks occurring in internal loci, whereupon the heterodimers, in conjunction with the Sir protein complex, localises to the break (Martin *et al.*, 1999). It is found to be involved in regulation of telomere length (telomere length maintenance - TLM) and silencing. TLM was assayed in mutant cell lines, using as probes genes naturally located at trypanosome sub-telomeric regions. Both transcriptionally active and inactive sub-telomeric regions were analysed in this fashion. This assay was used in the bloodstream form mutants. Telomere silencing was monitored also in procyclic form mutants, using RT-PCR to analyse the transcriptional status of three different characterised *MVSG* genes.

An assay was also developed to look at the effect disruption of the heterodimer had on the localisation of mini-chromosomes at the nuclear periphery. It was hypothesized that loss of Ku would result in mislocalisation of the mini-chromosomes in the nucleus, consistent with evidence that Ku is involved in tethering telomeres to the nuclear periphery. Localisation was monitored using fluorescent in situ hybridisation (FISH), while general minichromosome maintenance was assayed by separating the chromosomes on a pulsed field gel and looking for maintenance or loss of minichromosomes.

4.2 Growth of *ku* mutants

The first phenotype examined was growth. The population doubling time of wild-type 3174 trypanosomes (9.35 and 9.75 h in 2 experiments) compared well with those of the heterozygous *ku70* and *ku80* mutants 70P3.1(+/-) (10.25 h), 70P4.2(+/-) (10.25 h), 80B1.1(+/-) (10 h) and 80B1.3(+/-) (9.12 h), as well as those of the null mutants 70P3.1B(-/-) (9.75 h), 70P4.2B(-/-) (10 h), 80B1.1P(-/-) (9 h) and 80B1.3P(-/-) (9.25 h). Ku therefore is not essential for bloodstream trypanosome growth *in vitro* and its absence causes no detectable increase in population doubling time, indicating there is neither a general delay in the cell cycle nor chronically death of some of the population.

4.3 Telomere Length Maintenance (TLM) in *ku70* and *ku80* bloodstream form mutants

Telomere length maintenance depends on Ku in yeast. In *T. brucei*, it is possible to monitor the length of individual telomeres because of the subtelomeric location of individual *VSG* genes. By digesting DNA at a restriction enzyme site upstream, but not downstream, of a *VSG* and then probing for that gene, a fragment containing the entire telomere fragment, including the telomere tract can be detected. We have studied two distinct *VSG*s. The *221* gene is in the transcriptionally active telomere in the bloodstream trypanosomes under study. The *VO2* gene has one copy in a silent telomere, and a second copy which, being within a chromosome, is flanked on both sides by restriction sites. For *KU70*, the clones 70P3.1(+/-) and 70P4.2(+/-), and a number of subclones of 70P3.1B(-/-) and 70P4.2B(-/-) were examined. Similarly, for *KU80*, we studied the clones 80B1.1(+/-) and 80B1.3(-/-), and the homozygous deletion clones 80B1.1P(-/-) and 80B1.3P(-/-).

We initially looked at the effect absence of Ku had on the telomere immediately downstream of the transcriptionally active expression site, *221* (Fig 22A). Sub-clones were derived for each independent *ku* homozygous and *KU* heterozygous mutant. One sub-clone from each independent *ku80* mutant was used (Fig 22C), whereas a more extensive analysis was carried out for *KU70* (Fig 22B). In this, five sub-clones were derived from each independent homozygous mutant and the progenitor wild-type cell

line, and a total of five heterozygous sub-clones were also analysed. Terminal restriction fragment sizes were measured by digestion with *Eco*RI and Southern blot analysis. For homozygous mutants these were generally smaller than those seen for heterozygous or wild-type cell lines (Fig 22B and 22C), indicating an impairment of telomere maintenance or an increase in telomere degradation. These same clones (for both *KU70* and *KU80*) were maintained in culture for approximately 150 generations when the TRF size was once again monitored. This allowed the generation of a heterogenous population of telomeres and revealed the effects of maintenance over a prolonged number of divisions. A comparison between initial clonal analysis and those heterogenous telomeres allowed to develop over time shows no major apparent difference in telomere fragment size (Fig 23A and 23B). This suggests that if either putative *KU* homologue is important in maintaining telomeres, the effect of their deletion does not cause a gradual loss of in tract sequence.

In order to quantify the extent of telomere loss in *ku* mutants, and to be certain this phenotype was due to the loss of Ku, we re-introduced wild-type *KU* into its own locus using a co-transcribed bleomycin (BLE) resistance cassette (pCC101), and as a control targeted BLE to the tubulin locus (pRM450) (Fig 24). Four null mutant sub-clones (70P4.2B-/-3, 70P4.2B-/-5, 70P3.1B-/-4, 70P3.1B-/-5; lanes 13, 15, 19, 20 in Figure 22B) were chosen for this experiment. Analysis of the active expression site (once again using *Eco*RI to release terminal restriction fragments) showed that in each case re-expressors harboured longer telomere tracts than did the null mutants (Fig 25A). These were grown for an estimated 18 generations during selection for transformants and then 12 more generations in the absence of bleomycin, at which time they were analysed for telomere tract length. Taking the fragment sizes for all these *KU70* experiments, the transcriptionally active 221 telomere fragment was a mean 7.2 kb long (range 3-10 kb) in homozygous mutant trypanosomes and 10.3kb (range 9-15 kb) in the re-expressor trypanosomes. This fragment contains about 3 kb of non-telomere tract sequence. This blot was subsequently probed with radiolabelled *KU70* ORF sequence to confirm absence of ORF sequence in cells transformed with pRM450 (Fig 25B).

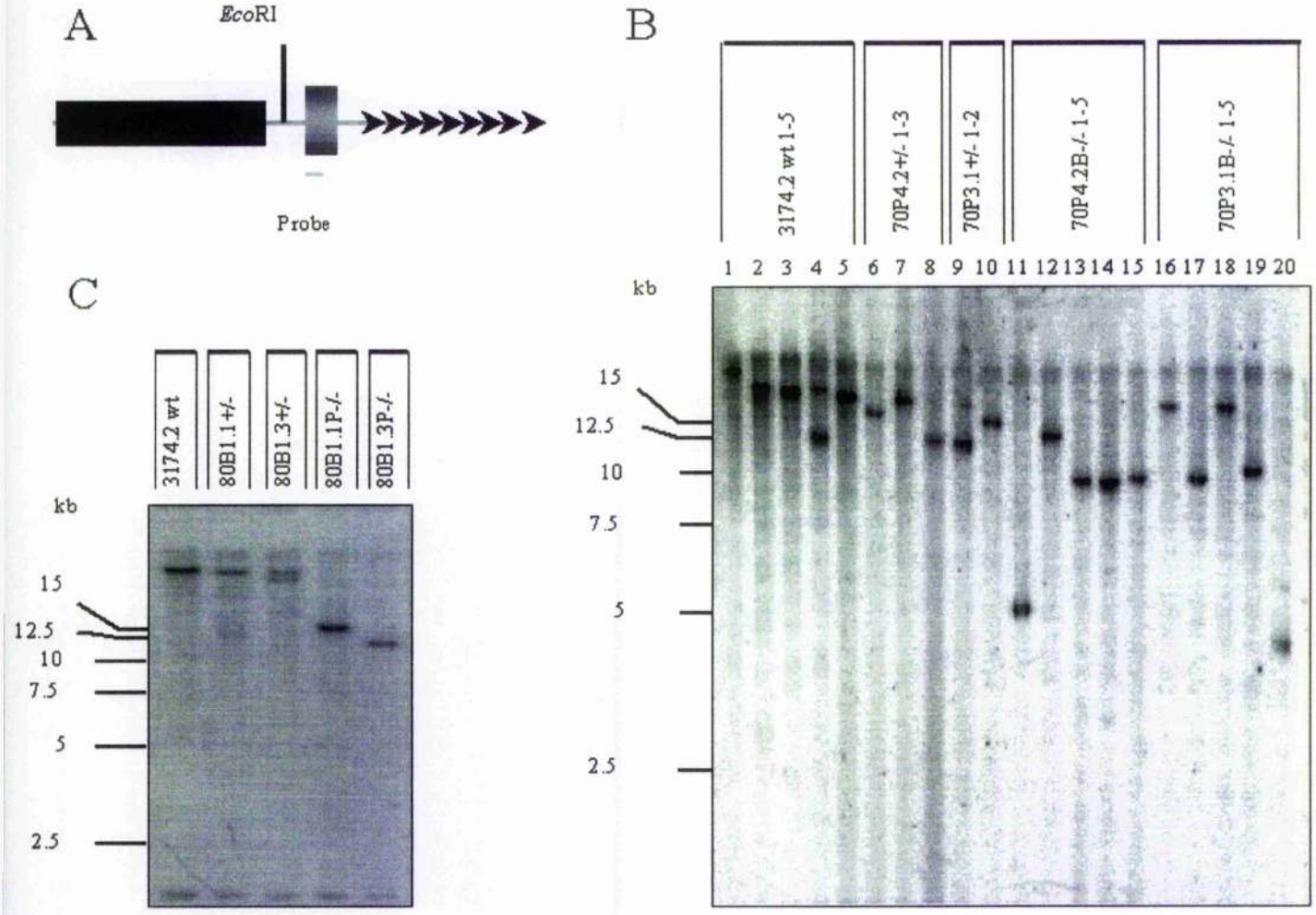

Figure 22. Telomere length maintenance in *T. brucei ku* mutants. (A) A brief map of the active expression site is depicted, with the telomere tract represented by a series of arrows, and an *Eco*RI restriction enzyme site upstream of *VSG221* gene (grey box) and the DNA sequence encoding the N-terminal region of *VSG221* used as a probe are shown. The 70-bp repeats are represented by a black box. (B) Genomic DNA for 20 sub-clones from wild-type (wt), heterozygotic (+/-) and homozygotic (-/-) *ku70* mutants digested with *Eco*RI, Southern blotted and probed with ^{32}P-labeled *VSG221* sequence. 3174.2 wt sub-clones 1-5 are in lanes 1-5 respectively; 70Pac4.2+/- sub-clones 1, 2 and 3 are in lanes 6, 7 and 8 respectively; 70Pac3.1+/- sub-clones 1 and 2 are in gel lanes 9 and 10 respectively; 70Pac4.2B-/- sub-clones 1-5 are in lanes 11-15 respectively; and 70Pac3.1B-/- sub-clones 1-5 are in lanes 16-20 respectively. (C) Genomic DNA from wt, *KU80* heterozygote (80B1.1+/-, 80B1.3+/-) and homozygotic mutants (80B1.1P-/-, 80B1.3P-/-) digested with *Eco*RI, Southern blotted and probed with ^{32}P-labelled *VSG221*.

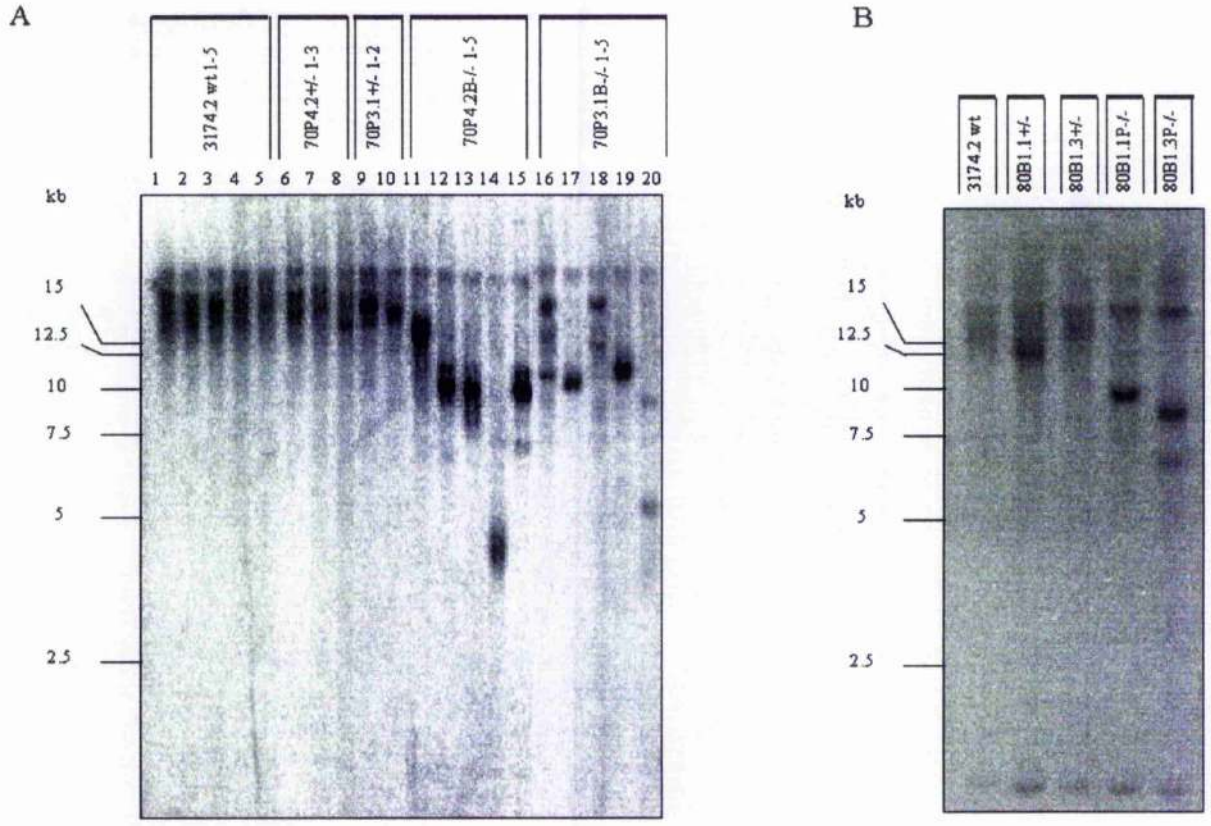

Figure 23. Telomere length maintenance in *T. brucei ku* mutants ~150 generations after cloning. (A) Genomic DNA for 20 sub-clones from wild-type (wt), heterozygotic (+/-) and homozygotic (-/-) *ku70* mutants after ~150 generations digested with *Eco*RI, Southern blotted and probed with ^{32}P-labeled *VSG221* sequence. 3174.2 wt sub-clones 1-5 are in lanes 1-5 respectively; 70Pac4.2+/- sub-clones 1,2 and 3 are in lanes 6, 7 and 8 respectively; 70Pac3.1+/- sub-clones 1 and 2 are in gel lanes 9 and 10 respectively; 70Pac4.2B-/- sub-clones 1-5 are in lanes 11-15 respectively; and 70Pac3.1B-/- sub-clones 1-5 are in lanes 16-20 respectively. (B) Genomic DNA from wt, *KU80* heterozygote (80B1.1+/-, 80B1.3+/-) and homozygotic mutants (80B1.1P-/-, 80B1.3P-/-) after ~150 generations digested with *Eco*RI, Southern blotted and probed with ^{32}P-labelled *VSG221*. Both blots were washed to a final stringency of 0.2 x SSC, 0.1% SDS at 65°C.

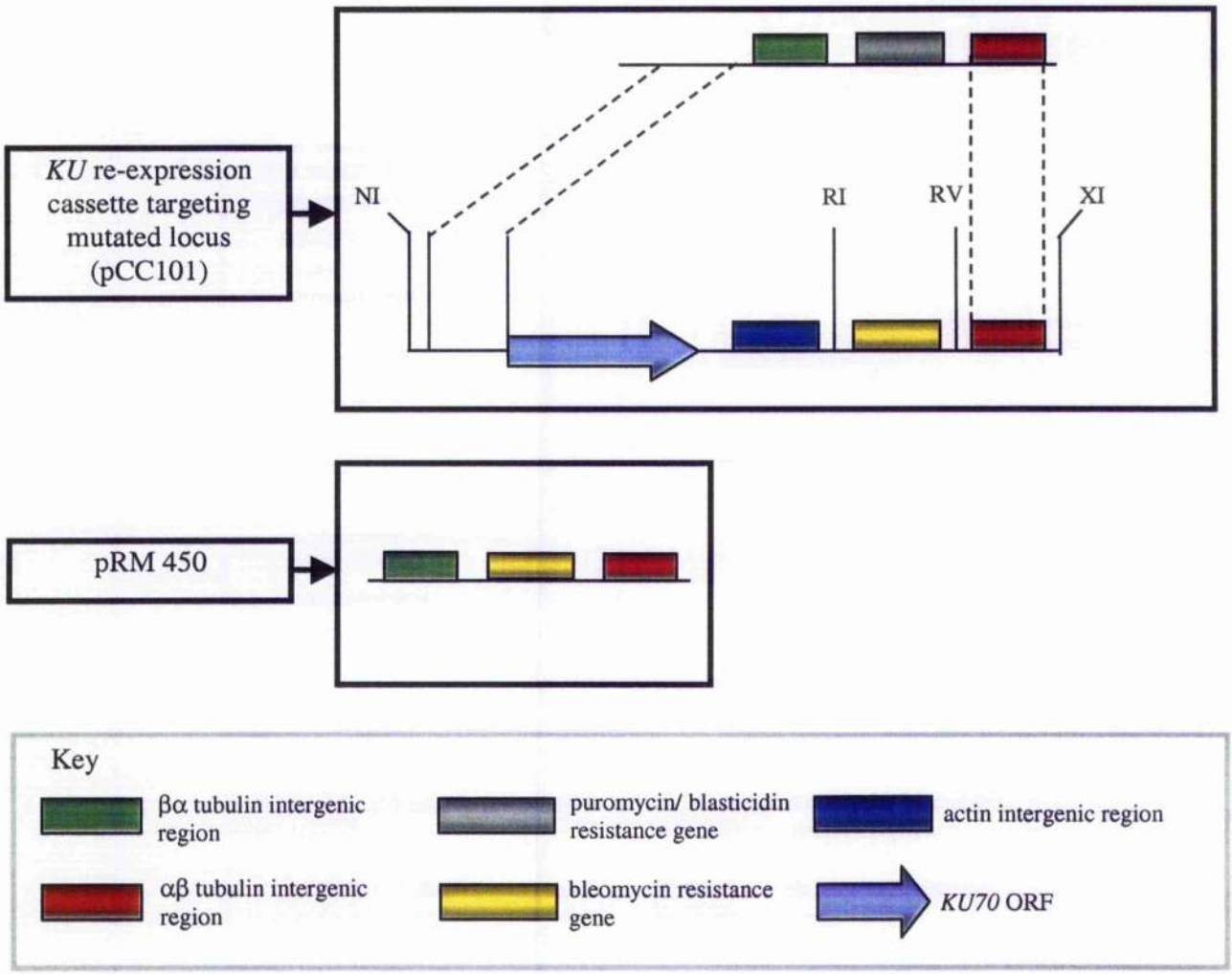

Figure 24. *KU* re-expression cassette (pCC101) and pRM 450 (electroporation control cassette). *KU70* was targeted to its original locus replacing an original disruption cassette. The 5' targeting flank used sequence immediately upstream of the ORF. Downstream targeting sequence utilised tubulin intergenic sequence originally used to provide processing signals for the selectable marker used to detect gene disruption. Bleomycin resistance was used to detect re-integration. pRM 450 targeted the tubulin array using tubulin intergenic sequences to allow for gene conversion. Bleomycin resistance was used for transformant selection. Abbreviations for restriction enzyme target sites are as follows: NI for *Not*I; RI for *Eco*RI; RV for *Eco*RV and XI for *Xho*I.

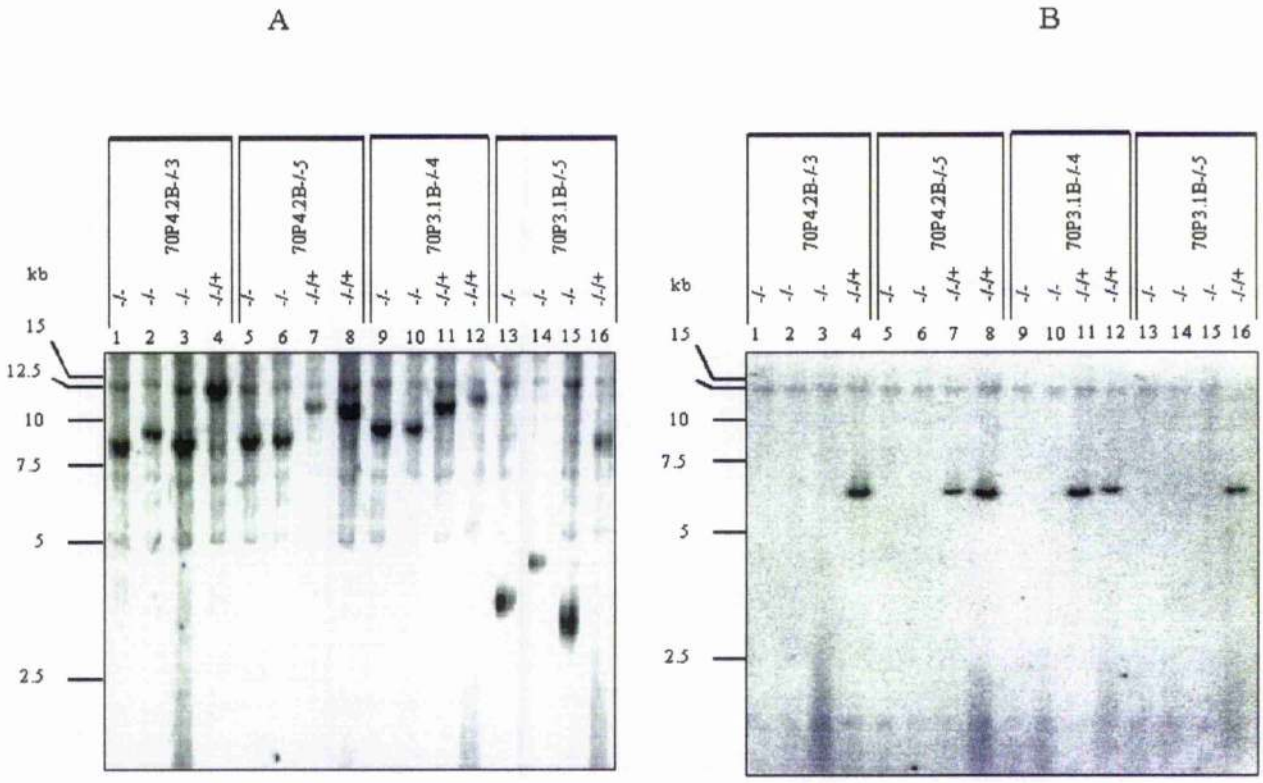

Figure 25. Re-expression of *KU70* in homozygous mutants and length of resultant active expression site TRFs. Genomic DNA was prepared from homozygous mutants transformed with *KU70* re-expression cassette and pRM 450. Genomic DNA for each cell line was digested with *Eco*RI, size fractionated on a 0.6% agarose gel and Southern blotted. The blot was probed with a ^{32}P radiolabelled probe containing DNA sequence encoding the N-terminus of VSG221. This was washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C (A). The blot was stripped with 0.1% SDS and subsequently probed with ^{32}P radiolabelled sequence belonging to the *KU70* ORF. Again the blot was washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C (B). Molecular markers (kb) are indicated to the left of the panel.

We next looked at the effect disrupting Ku had on a transcriptionally inactive telomere (Fig 26). This was carried out using restriction digestion with *Age*I and Southern blot

analysis. In this instance three of the four sub-clones used in the *221* experiment were analysed. The *VO2* gene has one copy in a silent telomere, and a second copy that, being within a chromosome, is flanked on both sides by restriction sites. Fragment sizes for the telomere downstream of *VO2*, measured on the same DNA samples and therefore of identical lineages as for the *221* analysis, are 9.4 kb (9.1-10 kb) in homozygous mutants and 9.8 kb (9.5-10.5 kb) in re-expressors. The telomere immediately downstream of the *VO2* gene has not been characterized, so we cannot say how much non-telomere sequence is present in the *Age*I terminal restriction fragments.

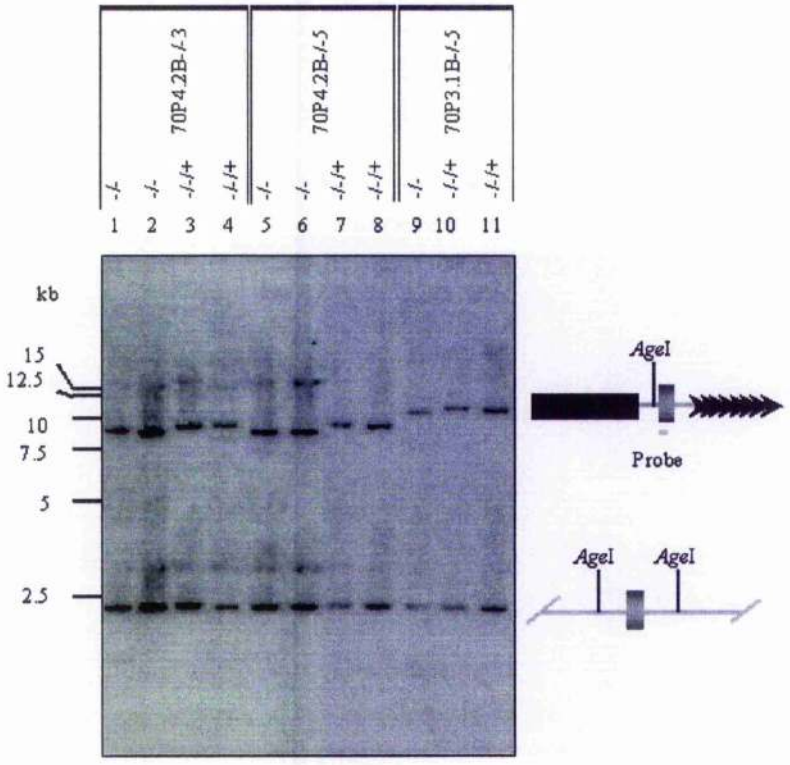

Figure 26. Telomere length maintenance of an inactive expression site in *T. brucei KU70* re-expressors. As in the analysis of telomere length in the active *VSG221* expression site (Figure 25), *T. brucei ku70* homozygous mutant clones re-expressing *KU70* are indicated by -/-/+, while clones depicted by -/- were transformed with the control plasmid pRM450. Cartoons to the right of the panel depict the *VO2 VSG* loci under investigation. The upper cartoon shows the telomeric *VO2* expression site, with telomere repeats depicted as arrows, *VSGVO2* gene as a grey box, and 70-bp repeats as a black box. An *Age*I restriction site present upstream of the telomeric *VO2* gene is shown, and the probe used for hybridisation depicted. The lower cartoon shows the chromosomal internal *VO2* gene copy as a grey box, and flanking *Age*I sites are shown. Genomic DNA from each transformed *KU70* sub-clone was digested with *Age*I, Southern blotted and probed with sequence encoding the N-terminus of VO2. Molecular size markers are indicated to the left of the panel.

4.4 Minichromosomal inheritance in *ku* mutants

Another proposed function of the Ku heterodimer is the tethering of telomeric DNA sequences to the nuclear periphery. Ku, along with other telomere binding proteins (e.g. Rap and Sir), acts as a physical intermediate between telomeres and nucleoporins (Galy *et al.*, 2000). The full relevance of this interaction is not entirely understood, but it is postulated to have a role in altering transcriptional activity of telomeric and sub-telomeric regions (Galy *et al.*, 2000). There are approximately 100 minichromosomes in the trypanosome genome ranging in size from 40-200 kb. They are believed to act solely as a reservoir for *VSG* genes and are transcriptionally silent. It was proposed that, in the absence of the docking protein Ku, minichromosomes would either be mis-localised in, or absent from, the trypanosome nucleus.

Initially to investigate whether minichromosomes were stably maintained in *ku* mutants, pulsed field gels (PFGs) were run to separate intermediate and minichromosomes. This experiment was originally carried out to analyse the *ku70* homozygous and heterozygous mutants. As for the Telomere Length Maintenance assay, a number of clones were chosen. Genomic DNA plugs were prepared for each clone and a PFG was used to separate the minichromosomes, which vary greatly in size due to heterogeneity. This procedure was carried out immediately after the cloning procedure and also ~150 generations after the cloning procedure. The separation is demonstrated in Figure 27 and displays no gross difference in ethidium bromide staining between the mutant and wild-type lanes. *ku80* homozygous and heterozygous mutants were analysed in a similar manner. Genomic DNA plugs were prepared for mutant and wild-type clones originally used for the TLM assay. As for the TLM assay, a single clone was utilised for each independent heterozygous and homozygous mutant cell line, each clone having been generated through an independent transformation. Chromosomal DNA was separated using identical conditions as for the *KU70* experiment (Fig 28A). Plugs were also prepared after these clones had been maintained in culture for approximately 150 generations and were separated using a PFG in an identical manner. This would help us elucidate whether minichromosomal loss, if occurring, was rapid or gradual and would

also enable visualisation of a more heterogenous population of cells. Again, stained gels showed no relative difference in minichromosomal DNA intensities between mutant and wild-type lanes. This was seen at the time of, and after, cell line cloning, suggesting at least no gross disappearance of mini-chromosomes in each cell line being analysed, with the intensity of the dominant minichromosomal bands being consistent between clonal counterparts (Fig 28).

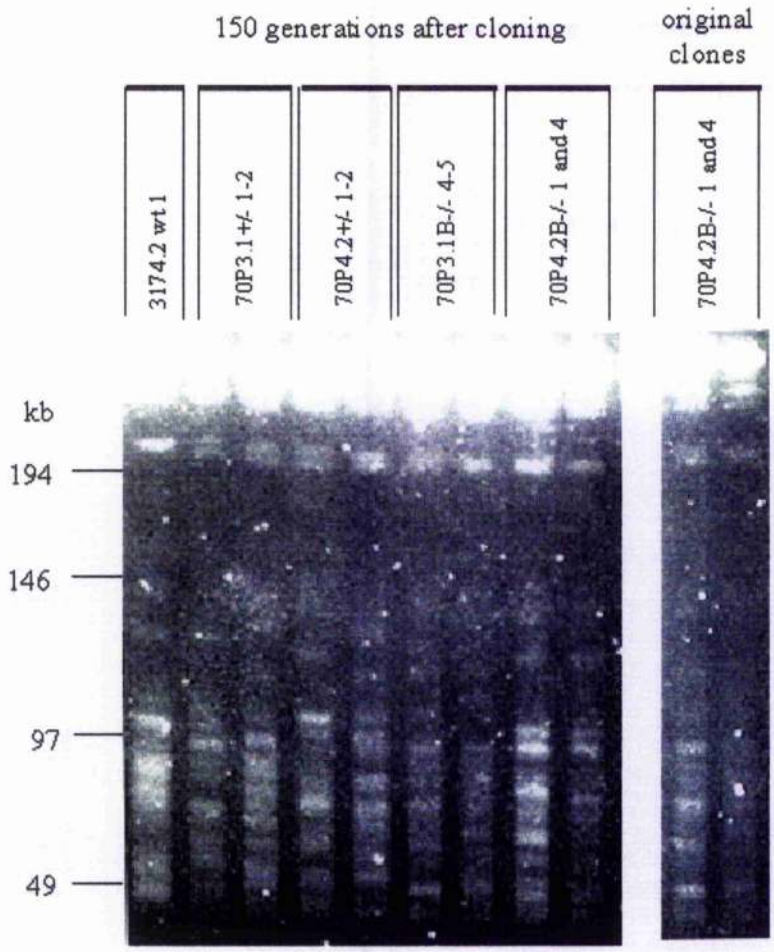

Figure 27. Mini-chromosome PFGs. Two independently transformed heterozygous and homozygous mutants (originally used for TLM assay) were cloned on semi-solid HMI-9 agarose dishes. 2 clones were prepared for each independent heterozygous and homozygous mutant aswell as a single wild-type clone. Each clone was maintained in culture for 150 generations and the mini-chromosomes were separated on a PFG. Genomic DNA plugs from original clones 70P4.2B-/- 1 and 4 were also separated on this gel.

Two minichromosomes of 55 and 60 kb have been found to have an rRNA gene promoter near one telomere, because of what is thought to be a recombination with a repetitive rRNA gene locus in a megabase chromosome (Zomerdijk *et al.*, 1992). When a drug resistance gene, *neo*^{r}, with the proper RNA processing signals was inserted downstream of this promoter in this 55-kb molecule, the resulting trypanosomes became drug resistant, indicating that the minichromosome is stably inherited during mitosis. The presence of this marker in two minichromosomes allowed a means to analyse individual minichromosome presence and size in *ku* mutants. Both PFGs were Southern blotted and probed with ^{32}P radiolabelled rDNA promoter sequence. The blots were then washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C (Fig 28). The results of this probing confirmed the presence of the marker in both mutant and wild-type genomes. Interestingly, the band hybridising to the probe sequence in both mutant lanes proved to be smaller or equal in size to those hybridising minichromosomes in the wild-type and heterozygous mutant lanes. Variation in band size between clones was believed to be due to heterogeneity in telomere length, homozygous mutants harbouring smaller marker minichromosomes probably due to decreased telomeric sequence.

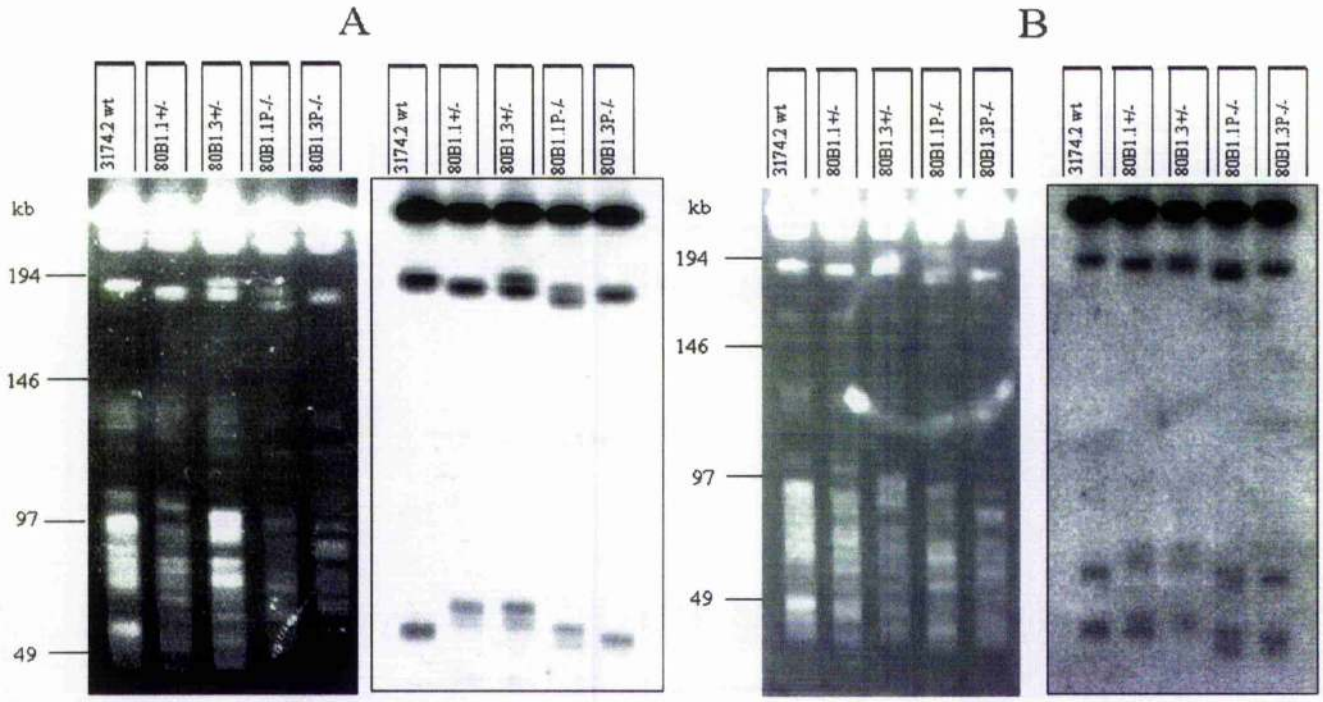

Figure 28. Maintenance of mini-chromosomes in *ku80* mutants. Two independently transformed heterozygous and homozygous mutants were cloned on semi-solid HMI-9 agarose dishes. Genomic DNA plugs were prepared for each clone and mini-chromosomes were separated on a Pulsed Field Gel (PFG). The gel was stained with ethidium bromide and Southern blotted. The blot was subsequently probed with ^{32}P radiolabelled ribosomal DNA promoter sequence (Panel A). Each clone was maintained in culture for 150 generations and the same conditions were used to separate the mini-chromosomes on a PFG. Again, the gel was Southern blotted and probed with the same DNA sequence (Panel B). Molecular markers (kb) are indicated to the left of each ethidium bromide gel.

4.5 Minichromosome and telomere localisation in *ku* mutants

Following the PFG minichromosome analysis, we deemed it necessary to analyse localisation of minichromosomes and telomeres in the trypanosome nucleus. Fluorescent in situ hybridisation (FISH) was used to visualise DNA segments for the individual segments under analysis. To locate minichromosomes, we used a 177-base pair (bp) repeat sequence as a probe. This sequence is specific for minichromosomes and has been used previously to study their distribution in interphase cells (Ersfeld and Gull, 1997). To visualise a DNA segment in telomeres, telomere repeats were used as a probe. This investigation was carried out in both procyclic and bloodstream form *ku* mutants.

It has previously been shown that the majority of nuclei contain 177 bp repeats clustered in less than half of the nucleus (Chung *et al.*, 1990). 177-bp repeats were either entirely clustered in a narrow band at one pole of the interphase nucleus or spread evenly in many small clusters restricted to roughly one half of the nucleus. In a less frequent pattern of distribution, the hybridising sequences were scattered throughout the entirety of the interphase nucleus in several small clusters (Chung *et al.*, 1990). In order to investigate localisation in *ku* mutants, both bloodstream and procyclic form mutants were examined. As expected, minichromosomes in both life-cycle stages in *KU* wild-type cells are predominantly located in a single half of the nucleus when viewed under a microscope and indeed in a number of cases are present as polar bodies (Figs 30 and 32). This appears to occur in both dividing and single cells. Bloodstream form mutants contain similar localisation patterns although clusters are more spatially divided in some of the cells under analysis (Fig 29). More singular bodies were observed in procyclic form wild-types and again are comparable with mutant cell lines (Fig 31).

The majority of telomeres are found clustered together, either extending in a network through the nuclear interior or localized at the nuclear periphery. In the majority of nuclei, the telomeric sequences are viewed under the microscope in several clusters spread throughout the entire nucleus or associated in several clusters that are uniquely located at the extreme periphery of the nucleus, indicating association with the nuclear envelope (Chung *et al.*, 1990). This was observed both in *ku* mutant and wild-type cells

upon analysis in both the bloodstream and procyclic form stages. Figure 33 demonstrates mutant lines probed with telomere sequence. Each of the three images in this figure contains clustered telomeric sequences at the nuclear periphery along with randomly spaced clusters. *KU* wild-type cells are shown in Figure 34 and display a similar pattern. Images one and three show more random clustering than seen in the other two images where predominant aggregation is found at the periphery. Procyclic mutant and wild-type cells probed with the same sequence also demonstrate asymmetric and polar distributions of telomeres (Fig 35 and 36). Interestingly, the nuclei of a dividing cell (Fig 35(3)) show patterns slightly different from one another, with the nucleus visible on the left-hand side displaying a more random arrangement. This nucleus does, however, contain some telomere localisation at the nuclear periphery.

These experiments suggest that Ku is not involved in important telomere tethering reactions in the nuclei of *T. brucei* cells. Ku is found to bind indirectly to nucleoporins in *S. cerevisiae* and it had been touted as a dominant player in docking telomeres to the periphery of the nucleus (Galy *et al.*, 2000). We additionally demonstrated that Ku is not involved in docking and localising the largely redundant minichromosomes in *T. brucei*. However, it must be realised that the same three-dimensional pattern can look very different when viewed under a microscope depending on the orientation and these results must be taken in that context.

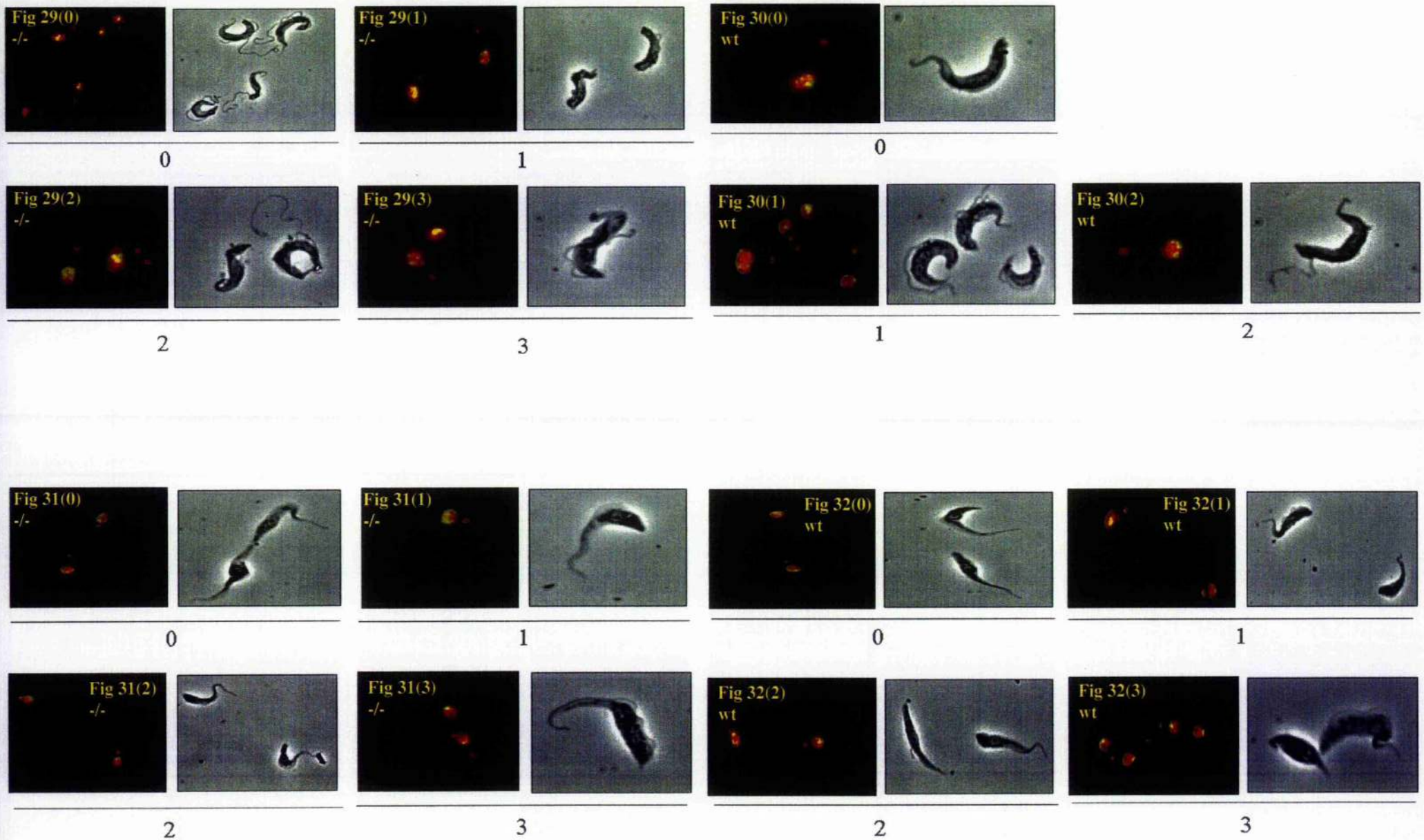

Figure 29-32. Minichromosomal localisation in *KU* wild-type and mutant cell lines. Fig 29 = bloodstream form *ku80* mutant cells; Fig 30 = bloodstream form *KU* wild-type cells; Fig 31 = procyclic form *ku70* mutant cells; Fig 32 = procyclic form wild-type cells. Right-handed images are phase optic images while left-handed images are fusions of DAPI and minichromosomal repeat probe stainings. Panel numbers represent document numbers of image files.

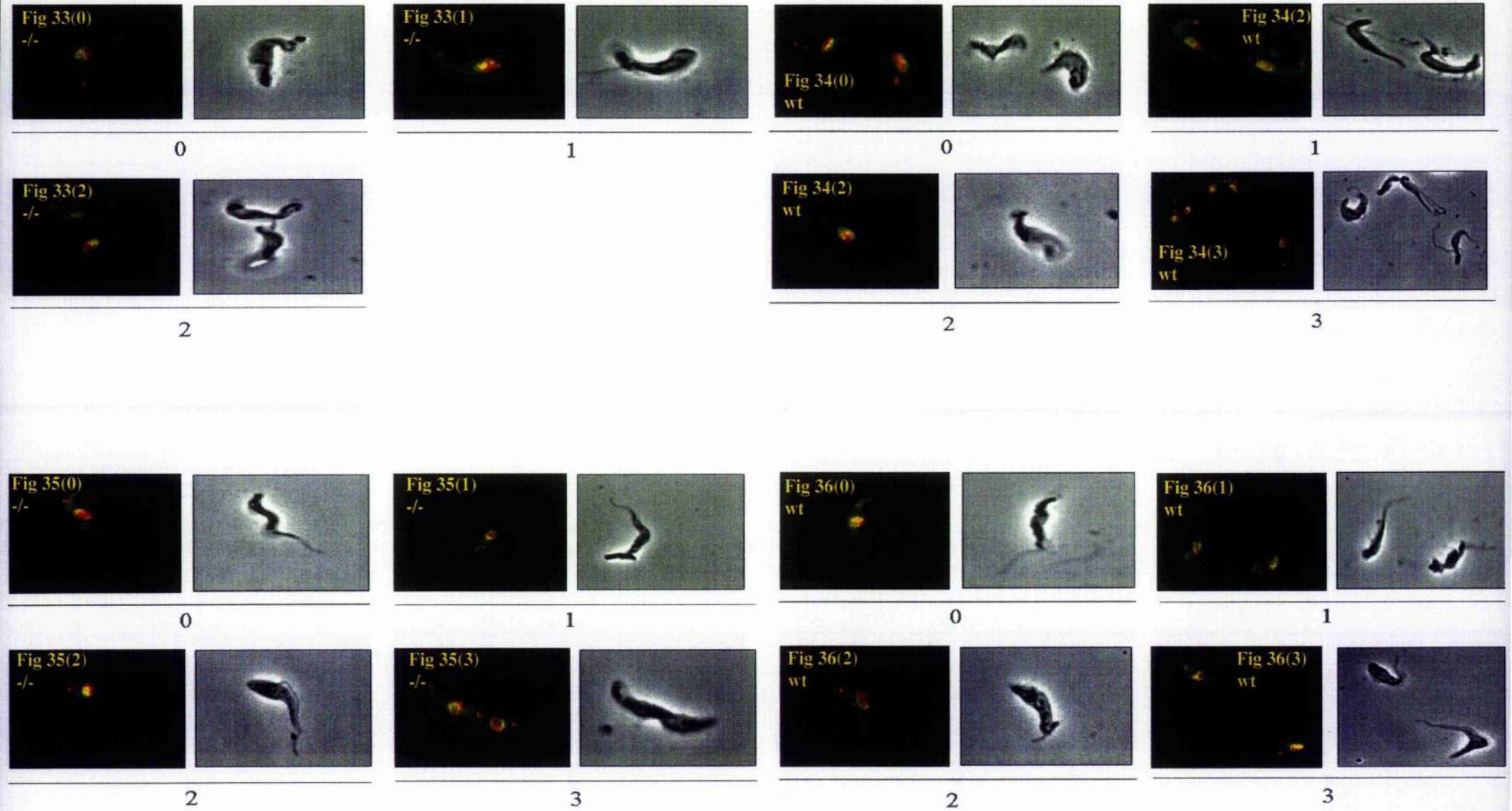

Figure 33-36. Telomere localisation in *KU* wild-type and mutant cell lines. Fig 33 = bloodstream form *ku80* mutant cells; Fig 34 = bloodstream form *KU* wild-type cells; Fig 35 = procyclic form *ku70* mutant cells; Fig 36 = procyclic form wild-type cells. Right-handed images are phase optic images while left-handed images are fusions of DAPI and telomere repeat probe stainings. Panel numbers represent document numbers of image files.

4.6 MMS sensitivity assay

Along with general telomere maintenance and localisation, Ku has been demonstrated to be involved in a form of DNA repair known as non-homologous end joining (NHEJ). Disruption of either gene causes a reduction in DNA double strand break repair via NHEJ. NHEJ is a very simple form of repair that generally involves a ligation reaction joining two DNA molecules on either side of a double strand break. Various assays have been utilised to look at this mechanism in both yeast and mammalian cells. However, unlike African trypanosomes, these cells have the ability to propagate circular plasmids, hence allowing a simple means to look at NHEJ. Plasmids can be digested with various restriction enzymes, the linear molecule is then transformed into cells and their capacity to perform NHEJ is measured as a function of the number of colonies having repaired and circularised the plasmid. An analysis of NHEJ in trypanosomes meant the development of a variety of assays (see Chapter 5), however to look crudely at how the mutagenesis of *KU* affected cells' ability to repair DNA double strand breaks, a partially radiomimetic drug methyl methane sulfonate (MMS), which causes point mutations both single- and double-strand breaks (Schwartz *et al.*, 1989), was added to HMI-9 medium and cell viability was subsequently monitored.

The same clones as used for the initial growth analysis were propagated in the continuous presence of MMS. Standard growth curves revealed a dose response effect on the growth of wild-type trypanosomes over the range 0.0003 – 0.0005% MMS (Fig 37A). The wild-type 3174.2 trypanosomes displayed population doubling times of 9.35 h in the absence of MMS, 13 h at 0.0003% MMS, 16.5 h at 0.0004% and 28 h at 0.0005%. This compared well with those of the heterozygous mutants where population doubling times were 9.56 h in the absence of MMS, 12.4 h at 0.0003%, 16.6 h at 0.0004% and 27.5 h at 0.0005%. Homozygous mutants displayed similar population doubling times: 9.13 h in the absence of MMS, 11.75 h at 0.0003%, 15 h at 0.0004% and 25.25 h at 0.0005%. These results represent an average of two independently obtained data sets. Thus, for both heterozygous and homozygous mutants of both *KU70* and *KU80*, the same dose response effect occurred, indicating no increase in mutagen sensitivity in the mutants.

Confirmation of the apparent lack of increased sensitivity in the mutants came from a clonal growth assay, in which 5 trypanosomes were plated in each well of a 96 well plate and the percentage of wells displaying growth after 96 h. This number was chosen after a series of pilot experiments were carried out to determine the number of cells that were required per well to ensure over 50% viability in the absence of MMS in the medium. This more sensitive assay showed, again, no increased sensitivity to MMS (Fig 37B).

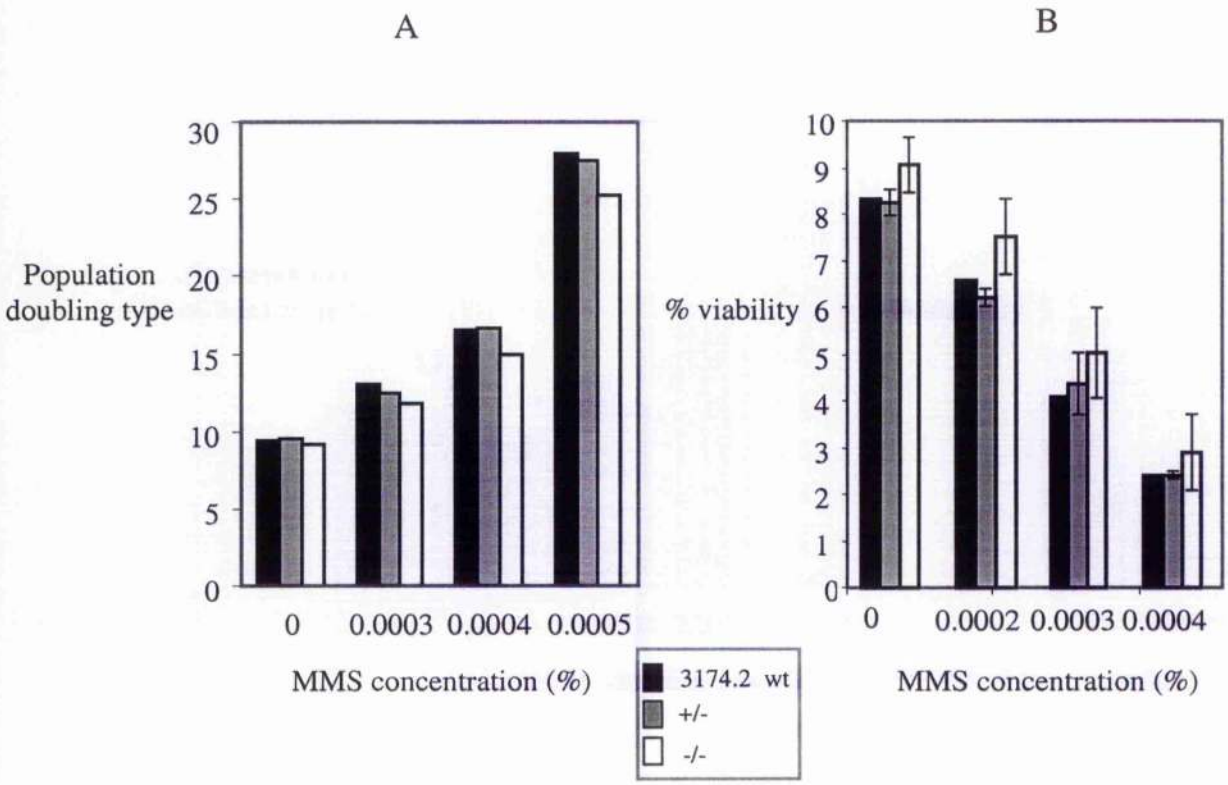

Figure 37. MMS sensitivity. (A) Growth of lines (*KU80* wild type, heterozygous and homozygous cells) was monitored over a specified period of time and population doubling times measured in medium containing increasing concentrations of MMS. (B) *KU70* MMS survivability assay. Two independent homozygous and heterozygous mutant clones and a single wild-type clone were tested for their sensitivity to MMS (methyl methanesulphonate). 5 cells were plated per well of a 96 well plate. Three independent MMS dilutions were carried out for each cell line; 0.0002%, 0.0003% and 0.0004% MMS. % viability was calculated as the % of wells to contain actively dividing trypanosomes 96 hours after plating.

4.7 *VSG* switching assay

Having established some of the general phenotypes predicted for *KU* genes, we next asked whether there is a specific role in the differential expression of *VSG* genes. In the bloodstream stage of *T. brucei* S427, switching occurs at a relatively low rate when compared with pleomorphic trypanosomes and is achieved by a number of routes, including transcriptional switching between telomeres and gene duplication into telomeres. Both *KU70* and *KU80* were mutated in a transgenic trypanosome strain, 3174.2, which had been engineered to assay for *VSG* switching (Rudenko *et al.*, 1998). The strain contains antibiotic resistance markers for both hygromycin and G418 in its active expression site. Growth of cells on both hygromycin and G418 allows the selection of cells expressing only the modified expression site and therefore the VSG coat encoded by the *VSG221* gene. Mice can be immunised against the VSG coat being expressed from the active expression site and by injecting those same mice with a known number of trypanosomes, the number of cells evading the immune system can be estimated and their switching frequencies measured. 24 h after injection, surviving trypanosomes, those which had switched their VSG coat and so were not killed by immune lysis, were recovered from the mice and plated over 96-well plates in HMI-9. After 10 days the number of wells which showed growth were counted, and the frequency of VSG switching was estimated. Furthermore the drug sensitivities of the cells from the wells showing growth were assessed, and genomic DNA was isolated from the switched trypanosomes to allow PCR amplification of marker genes from the modified expression site. Using this system, three types of VSG switching event can be distinguished which result in the expression of a novel *VSG* gene. Cells which are sensitive to both hygromycin and G418 may have undergone either an *in situ* transcriptional switch or a long-range expression site gene conversion. PCR amplification using primers directed against the hygromycin phosphotransferase (Hyg), neomycin phosphotransferase (Neo) and *VSG221* (221) ORFs differentiates between these possibilities. During an *in situ* switch, the active expression site is transcriptionally silenced, while another expression site is activated, leaving all markers intact and allowing amplification of the Hyg, Neo and 221 products (Hyg+, Neo+ and 221+). These events could also be reciprocal

exchanges upstream of the hygromycin marker, but in practise transcriptional switching is most common in this strain. In an expression site gene conversion all the markers are replaced by sequence from another expression site so they cannot be amplified by PCR (Hyg-, Neo- and 221-). Cells which have undergone the third type of switching event, *VSG* gene conversion, are resistant to hygromycin but sensitive to G418. *VSG* gene conversion involves the replacement of the sequence from the 70-bp repeats to the 3' end of VSG221 with sequence from another expression site, removing the Neo and 221 markers but leaving the Hyg marker intact (Hyg+, Neo- and 221-).

This assay was carried out for *ku70* mutants with two independently derived homozygous (70P3.1B(-/-) and 70P4.2B(-/-)) and heterozygous mutants (70P3.1(+/-), 70P4.2(+/-)) being used. All gave the rate of $0.1 - 0.6 \times 10^{-6}$ switch/cell/generation (Fig 38), which is typical of wild-type cells. Thus, there is no evidence for a significant decrease in switch rate in the absence of *KU70*.

Further, the relative levels of transcriptional versus recombinational *VSG* switching were measured. Examination of the switching events from a selection of the switched variants recovered indicated that there was no gross change in the relative levels of transcriptional versus recombinational *VSG* switching (Table 6).

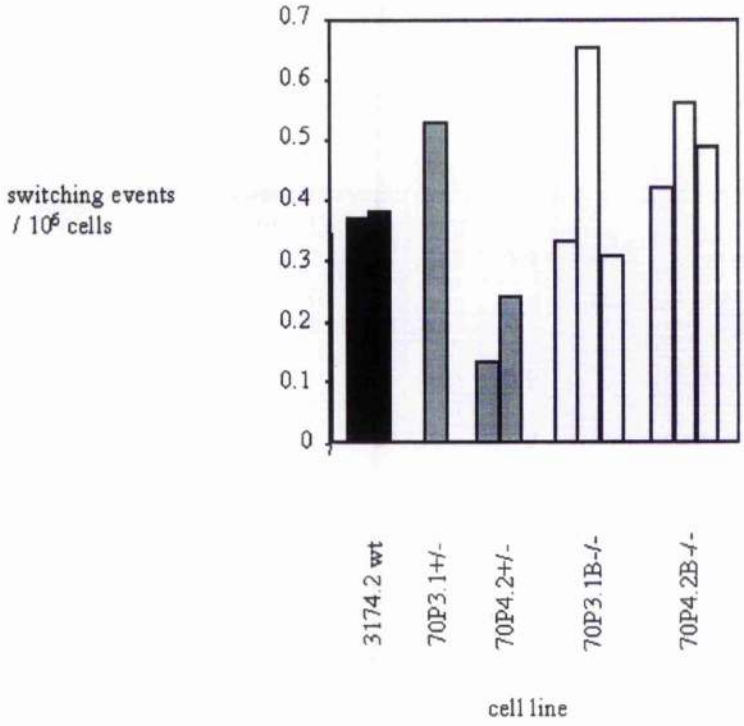

Figure 38. *VSG* switching frequencies in *ku70* mutants. Wild-type cells (3174.2 wt) were compared with two *KU70* heterozygous mutants (70P3.1+/-, 70P4.2+/-) and two homozygous mutants (70P3.1B-/-, 70P4.2B-/-). Each bar in this graph represents an independent experiment.

Cell line (No. of switched variants analysed)	**Gene conversion expression site**	**Gene conversion VSG**	***In situ* switch**
3174 (20) 1 mouse in duplicate	25%	50%	25%
***KU70+/-* amalgamation (21) 3 mice**	30%	47%	23%
***ku70-/-* amalgamation (59) 4 mice**	28%	38.5%	33.5%

Table 6. *VSG* switching frequencies of *ku* mutants. PCR amplification and antibiotic sensitivities were used to analyse the type of switching reactions used by cells to switch their surface coat.

4.8 Ku is not essential for silencing of the naturally telomeric *MVSG* genes

As mentioned in the introduction, it is plausible that Ku regulates differential expression of *VSG*s in the metacyclic stage of *T. brucei*. In that stage, a subset of telomeric genes known as *MVSG*s becomes differentially activated at the transcriptional level, generating a population diverse for VSGs. As these are the only trypanosome protein-coding genes known to have their own promoters, and to be subject to absolute control by transcription initiation, it has been speculated that they have achieved such control through operation of a TPE on the *MVSG* promoters, which are only ~5 kb from the telomere tract (Borst *et al.*, 1998).

VSG genes are first expressed in the metacyclic stage of the trypanosome life cycle, in the tsetse. Differential expression of the subset of telomeric genes used there occurs by transcriptional activation. In most life cycle stages, *MVSG*s are not expressed, but individual ones are activated in individual metacyclic trypanosomes. Here, the hypothesis that silencing of *MVSG*s in the procyclic (tsetse midgut) stage is exerted by a Ku-mediated telomere position effect was tested. We have tried repeatedly to transmit *T. brucei* S427 through tsetse, to no avail (J.D. Barry, data not shown), so have been unable to identify MVSGs in this strain. Therefore, we instead used *T. brucei* EATRO 795, which routinely develops to the metacyclic stage, permitting characterisation of the *MVSG*s encoding the ILTat 1.22, 1.61 and 1.63 VSGs. We deleted both *KU70* alleles in procyclic stage *T. brucei* EATRO 795 grown *in vitro* (see Chapter 3), creating the independent heterozygous, puromycin-resistant knockout cloned lines 70+/-1.1 and 70+/-1.2, and then their homozygous, blasticidin-resistant doubly deleted clonal descendants called, respectively, 70-/-1.1(6) and 70-/-1.2(1). RT-PCR was used to search for transcripts from two regions of each *MVSG* locus. The first, downstream of the transcription start sites (Fig 39), would detect primary transcripts. The second, including the amino-terminal VSG coding sequence, would detect primary transcripts and stable RNA. The scant availability of metacyclic stage RNA prevented its use as a positive control, so instead of metacyclic cDNA we used the corresponding trypanosome genomic DNA. Although this control (lanes 11 in Fig 39A, 39B), and an RNA POLI positive

control for integrity of cDNA, were both positive (Fig 39C), no RT-PCR products were detected in wild-type, heterozygous or homozygous mutants for any of the three *MVSG* loci immediately downstream of their promoters (Fig 39A). Similarly, nothing was detected for the coding region (Fig 39B).

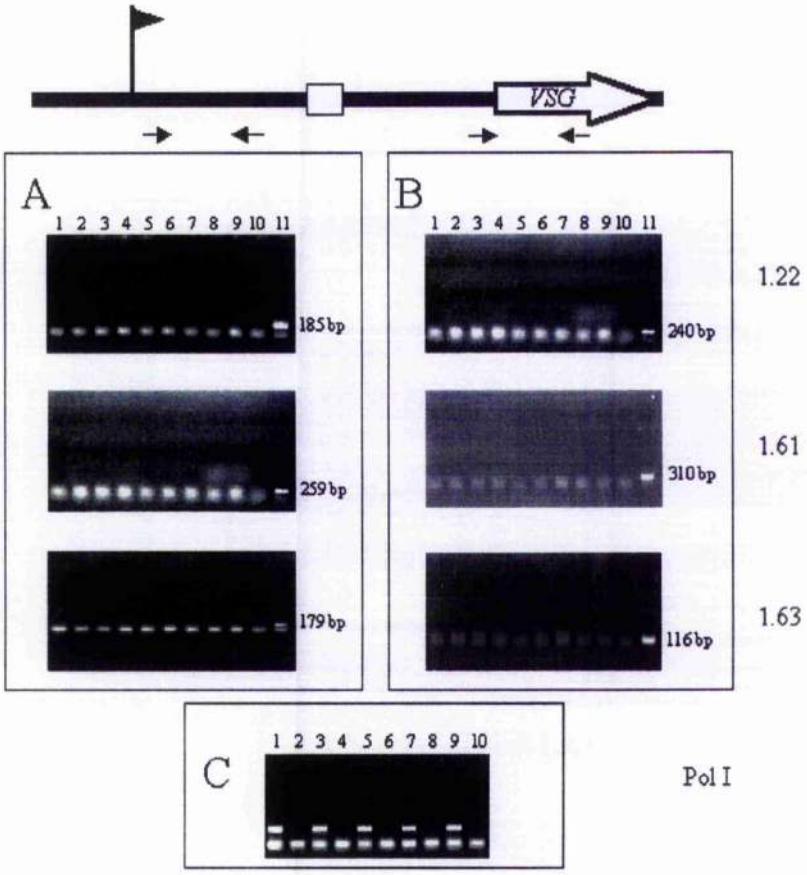

Figure 39. Transcriptional status of telomeric *MVSG*s in procyclic form *ku70* mutants. cDNA (RT+) and mock cDNA (made without reverse transcriptase; RT-) was generated from total RNA for the following cell lines: *KU70* wild-type (EATRO 795), *KU70* heterozygotic mutants 1.1+/-(6) and 1.2+/-(1) and *ku70* homozygotic mutants 1.1-/-(6) and 1.2-/-(1). The telomeric environment of the three *MVSG*s (1.22, 1.61 and 1.63) examined is depicted in the top diagram: the promoter is shown as a black flag, the small number of 70 bp repeats as a white box and the *VSG* gene a white arrow; below the expression site arrows show the positions of oligonucleotides used for PCR. Panel A shows PCR carried out on the cDNA and mock cDNA for each *MVSG* proximal to the promoter, while panel B shows PCR of the three *MVSG* sequences. In each panel, lanes 1 (RT+) and 2 (RT-) shown are for wild type cDNAs, lanes 3 (RT+) and 4 (RT-) from 70+/-1.1, lanes 5 (RT+) and 6(RT-) from 70+/-1.2, lanes 7 (RT+) and 8 (RT-) are from 70-/-1.1(6), lanes 9 and 10 are from 70-/-1.2(1) and lane 11 is a control PCR using the respective primers on EATRO 795 procyclic genomic DNA. Panel C shows RT- PCR amplification of the RNA polymerase I large subunit as a control for cDNA integrity; lane order is identical to panels A and B, although no genomic DNA control was performed.

CHAPTER 5

DNA REPAIR PATHWAYS IN *T. brucei*

5.1 Introduction

DNA repair is imperative for the successful maintenance and propagation of genetic information. DNA double-strand breaks (DSBs) can occur spontaneously, due to replication errors, or may be induced otherwise by exogenous agents. There are two principal forms of DSB repair in eukaryotes and they differ in their requirement for DNA homology. Homologous DNA sequence is essential for DNA DSB repair by homologous recombination and this is the primary mechanism of DSB repair in yeast species. Vertebrates are believed to use preferentially the second type of repair system, termed non-homologous end joining (NHEJ). NHEJ is a less precise mechanism and is largely independent of terminal DNA sequence homology, producing junctions that can vary in their sequence composition.

Participation of NHEJ in DNA DSB repair has never been demonstrated formally in African trypanosomes. To investigate its occurrence in *T. brucei*, a number of assays were developed mainly using exogenous DNA sources as substrates for DNA repair systems inherent in the trypanosome. Initial experiments looked at the fate of exogenously introduced restriction enzyme target sites after transient transformation with cassettes encoding the restriction enzyme. A second assay looked at restoration of bacterial transformation efficiency (i.e. recircularisation) of restriction enzyme digested plasmids following electroporation into trypanosomes. This assay is used standardly in yeast to test for the presence and efficiency of DNA repair by illegitimate recombination reactions (Milne *et al.*, 1996). This is possible due to the ability of yeast cells to propagate plasmid DNA. Our assay was essentially a "shuttle" assay, in so far as trypanosomes do not naturally facilitate plasmid replication. DNA was rescued from recovering *T. brucei* cells after electroporation and this DNA was used in an attempt to transform bacteria, with bacterial transformants representing plasmids repaired by the trypanosome population under analysis.

A final analysis looked for the presence of NHEJ in homologous recombination-deficient trypanosomes. Disrupting this element of DNA repair would hopefully lead to other forms of repair becoming detectable, and even up-regulated. Rad51, in yeast a

member of the Rad52 epistasis group (integral in yeast homologous recombination), had previously been demonstrated to be involved in DNA repair in trypanosomes (McCulloch & Barry, 1999). *rad51* mutants were electroporated with cassettes containing non-compatible ends that would prevent their integration into the endogenous genome via conventional homologous recombination. This cassette also contained promoter DNA sequence to allow selection in the event of integration into non-transcribed regions of the genome. Study of the junctions encompassing the integration sites of the cassette would subsequently allow investigation into how the cassettes were actually integrated and tell us whether NHEJ was used, and if so, how efficient the process is in trypanosomes.

5.2 "Shuttle" plasmid assay

The first approach taken to assess the ability of *T. brucei* to repair DNA DSBs by NHEJ used linear plasmids transiently to transform trypanosomes. Re-circularisation of plasmid DNA by repair machinery could only be due to illegitimate recombination due to the fact that trypanosomes lack internal substrates to repair pBluescript by conventional homologous recombination reactions.

Plasmid DNA (pBluescript) was digested with *Eco*RV and the digest was used to transform trypanosomes (Fig 40). 3.5 μg of digested DNA was used per trypanosome transformation, with repair capacities of both procyclic and bloodstream form cells being analysed. Trypanosome strains EATRO 795 (procyclic) and ILTat 1.2 (bloodstream) were used in this experiment. Transformants were allowed to recover for 24 h with ~5.0 x 10^7 procyclic and ~4.0 x 10^7 bloodstream form cells being harvested for DNA. 600 ng of DNA extracts (extracted using genomic DNA preparation; see section 2.2.2) were then used to transform supercompetent bacteria and transformants were selected as colonies on L-agar plates containing ampicillin (0.27 M). Digested plasmid extract was tested for residual supercoiled DNA by gel electrophoresis of purified digested material. Transformants were counted for each individual bacterial transformation. Not surprisingly, transformation of the supercompetent cells with native wild-type trypanosome genomic DNA gave rise to no bacterial transformants (Table 7 – row A). This essentially assured us that there was no bacterial origin of replication along with an ampicillin resistance-encoding gene inherent within the trypanosome genome. When 600 ng of pBluescript was digested with *Eco*RV (Table 7 – row E), purified and used to transform supercompetent bacteria the result was 2400 bacterial transformants. This implied that the method of restriction enzyme digestion/purification was not sufficient to rid the preparation of supercoiled plasmid DNA. Bacterial transformations with DNA extracted from trypanosome transformants resulted in massive decreases in transformant numbers (>13 fold difference – Table 7, rows B and C). Indeed, there was no detection of bacterial transformants following transformation with DNA extracted from ILTat 1.2 bloodstream form transformants (Table 7 - row D). The results were inconclusive, for several possible reasons.

DNA	Colony numbers
(A) gDNA from EATRO 795 procyclic trypanosomes used to transform supercompetent bacteria (XL-1 blue MRF – Stratagene)	0
(B) EATRO 795 experimental transformation 1 (i.e. electroporated with purified plasmid digest)	180
(C) EATRO 795 transformation 2	25
(D) ILTat 1.2 transformation 1	0
(E) pBluescript digested with *Eco*RV and used to transform supercompetent bacteria (XL-1 blue MRF – Stratagene)	2400

Table 7. "Shuttle" plasmid assay 1

One possibility, as mentioned above, is that the DNA used in this experiment was not wholly digested, with maintenance of some supercoiled plasmid DNA in the preparation used to transform trypanosome populations. There appeared no evidence for remnant supercoiled plasmid DNA after size-fractionation of digested DNA on a 0.7% agarose gel. However, immediate transformation of bacteria with apparently digested DNA shows that traces of the supercoiled plasmid remain in the purified digest preparation. To circumvent this problem, a new method of plasmid digest purification was utilised.

pBluescript plasmid DNA preparations were once again digested with *Eco*RV. This time, however, plasmid DNA digests were size-fractionated on a 0.7% agarose gel and EtBr stained. The 3 kb DNA band representing *Eco*RV digested plasmid was gel extracted. Gel extracted product was subsequently butanol extracted to rid the sample of traces of EtBr. As with the original experiment, 3.5 μg of plasmid was used to transform the trypanosome cell line. Experimental transformations were carried out following

addition of DNA to prepared trypanosome populations prior to electroporation. Mock transformations were also carried out whereby digested and supercoiled pBluescript were added independently to trypanosome preparations after the electroporation had been carried out. As a further control, pBluescript (linear and digested) was added to SDM-79 medium containing 4 x 10^7 cells and DNA was extracted from this population and used to transform bacteria. This would inform us whether electroporation of plasmid had any effect on plasmid DNA preparation from trypanosomes exposed to pBluescript.

DNA	**Colony numbers**
(A) gDNA from EATRO 795 procyclic trypanosomes	0
(B) DNA from EATRO 795 electroporated with linear plasmid (1)	1
(C) DNA from EATRO 795 electroporated with linear plasmid (2)	0
(D) DNA from EATRO 795 electroporated with supercoiled DNA	>10,000
(E) DNA from linear mock transformation (EATRO 795)	0
(F) DNA from supercoiled mock transformation (EATRO 795)	33
(G) DNA extracted from trypanosomes to which linear DNA was added without trypanosome electroporation	1
(H) DNA extracted from trypanosomes to which supercoiled DNA was added without trypanosome electroporation	8
(I) pBluescript digested with *Eco*RV used to transform bacteria	32

Table 8. "Shuttle" plasmid assay 2

It is clear from this experiment (Table 8) that fractionation and purification of linearised plasmid via gel extraction is a far superior method of ensuring removal of supercoiled plasmid prior to bacterial or trypanosome transformation. There is more than a seventy-fold reduction in the number of bacterial colonies transformed to contain ampicillin resistance after heat-shock with linearised plasmid (Table 7 – row E, Table 8 – row I). It

is also apparent that the method used to extract DNA from trypanosomes transformed with pBluescript (supercoiled control) is sufficient for harvesting plasmid DNA prior to bacterial transformation (Table 8 – row D). The numbers of bacterial colonies resulting from transformations with DNA extracted from trypanosomes electroporated with linear pBluescript were also greatly reduced when compared with the previous experiment (>25 fold reduction, Table 8 – rows B and C). Results of the supercoiled mock trypanosome transformations (Table 8 - row F), and the transformation of bacteria with DNA preparations from trypanosome populations to which supercoiled plasmid had been added but not electroporated (Table 8 – row H), suggest that there is a low background level of supercoiled plasmid DNA that is maintained through DNA extraction. However, electroporation with the same plasmid structure (supercoiled) suggests that plasmid can efficiently transform cells and if NHEJ were a major form of DNA repair, DNA extracted from trypanosomes electroporated with linear pBluescript would likely produce a larger number of bacterial colonies upon transformation.

These results do not indicate the presence of DNA double-strand break repair by NHEJ in trypanosomes. This may, however, be for a number of reasons. Firstly, it could be the case that crude repair of exogenous DNA substrates is not a capability of trypanosomes. Secondly, gel extracted DNA may not be the ideal substrate for repair. Finally, the transformation efficiency of linear pBluescript may be too low to allow repair detection.

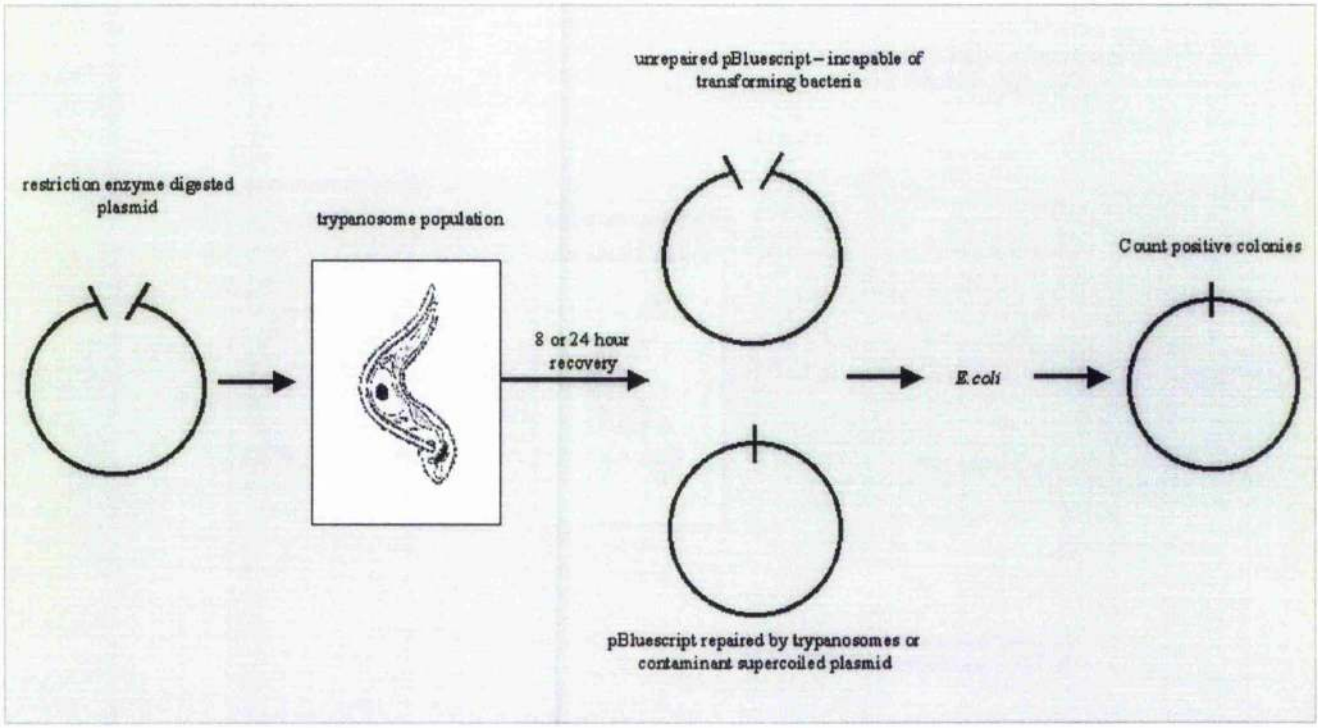

Figure 40. "Shuttle" plasmid assay cartoon. A fixed number of trypanosomes were electroporated with restriction enzyme digested plasmid. After a specified period of recovery, DNA was harvested from the trypanosome population. This DNA was then used to transform XL-1 blue MRF supercompetent bacterial cells. Bacterial transformants should only result from the repair of plasmid DNA and would allow calculations of trypanosome repair efficiencies.

5.3 Detection of NHEJ using HO endonuclease

A second assay used to monitor endogenous repair systems in the trypanosome involved expression of a enzyme, where the enzyme cleaves at a unique target site in the trypanosome genome. The initial aim of the experiment was the detection of DSBs following transient transformation with a cassette expressing the endonuclease. The occurrence of DSBs following transformation would subsequently allow analysis of how the break was repaired. Integration of the endonuclease target site into the tubulin gene array meant that induced breaks could be repaired by legitimate or illegitimate recombination reactions.

HO endonuclease, naturally found in yeast species, was the enzyme chosen for this experiment. Yeast strains are capable of switching mating type and the HO gene product catalyses the initial step in this process. The gene ORF encoding HO endonuclease was amplified using *Pfu* polymerase. This reaction used a 5 min denaturation step followed by 30 cycles of 1 min at 95°C, 1 min at 55°C and 5 min at 72°C. There was a final extension time of 10 min at 72°C. The template used for this reaction was 30 ng of the plasmid pGAL-HO (contains the *HO* ORF). The amplified product was purified and subsequently digested with *Pvu*II. Digestion product was then purified and quantified. pMLGGFP (donated by M.L. Ginger) was digested with *Sal*I and *Eco*RV and the GFP fusion protein was released. Overhanging ends produced as a result of restriction enzyme digestion were filled in using Klenow DNA polI and the plasmid (containing pBluescript and tubulin processing signals) was separated from the DNA fragment containing the fusion protein ORF by size fractionation using a 0.7% agarose gel. The plasmid was then gel extracted and quantified. The HO endonuclease gene ORF fragment was cloned into this plasmid, permitting mRNA production due to the presence of tubulin processing signals. Transcription was driven from a ribosomal DNA promoter (Fig 41A – pCC151).

Before the transient transformation experiments could proceed, introduction of the HO endonuclease target site to the trypanosome genome was necessary. It was decided to introduce this site to the tubulin array. This was performed using a cassette (pCC201, donated by R. McCulloch) containing tubulin targeting flanks surrounding a selectable marker and the endonuclease target site (Fig 41B). Prior to trypanosome transformation,

the cassette was digested with NotI and XhoI and purified. 5 μg of purified digestion product was used to transform 3 x 10^7 S427 procyclic trypanosomes and the transformants were selected using limited dilutions containing wild-type cells as feeders. Selection was carried out using hygromycin at a final concentration of 5 $\mu g.ml^{-1}$. Correct integration was detected by PCR analysis (see below).

5 μg of pCC151 DNA was used to transiently transform 3 x 10^7 S427 procyclic trypanosomes already engineered to contain the HO endonuclease target site. Two independent transformations were performed. One mock transformation, using no DNA for electroporation, was also undertaken in conjunction with the transformation driving transcription of the endonuclease. Transformed cells were allowed to recover at 27°C for 24 h prior to genomic DNA extraction. DNA was extracted from transformed and wild-type cells, digested with *Eco*RV, gel electrophoresed (0.7% agarose gel) and Southern blotted. The membrane was probed with α-^{32}P radiolabelled hygromycin resistance gene ORF sequence and was then washed to a final stringency of 0.2 x SSC, 0.1% SDS at 65°C. Successful transient transformation of the pCC151 should result in *in vivo* digestion of the *Eco*RV fragment harbouring the *HO* endonuclease target site in the tubulin gene array. This should theoretically, assuming repair is not very rapid, result in the production of two bands following the transient transformation of the plasmid pCC151. Alternatively, failure of transient transformation will result in the maintenance of a single band upon hybridisation of the blot with the radiolabelled hygromycin resistance gene ORF sequence.

It is clear from Figure 42 that several other hybridising bands are apparent along with the endogenous genomic DNA band. These bands were clearly visible as DNA fragments on the agarose gel used to separate out the genomic DNA digests. They appear to represent super-coiled and nicked forms of plasmid DNA (hybridisation cross-contamination; see Fig 42) that was initially used to transform trypanosomes containing the restriction enzyme target site. Apart from hybridising plasmid DNA, only a single band hybridises to the radiolabelled probe in both the mock control and transient transformation lanes, suggesting either that expression of the *HO* endonuclease has not occurred, that it has occurred at such a low level that digest products are not visible from

Southern blot analysis or that HO does not function in trypanosomes. Alternatively, repair may have been very rapid.

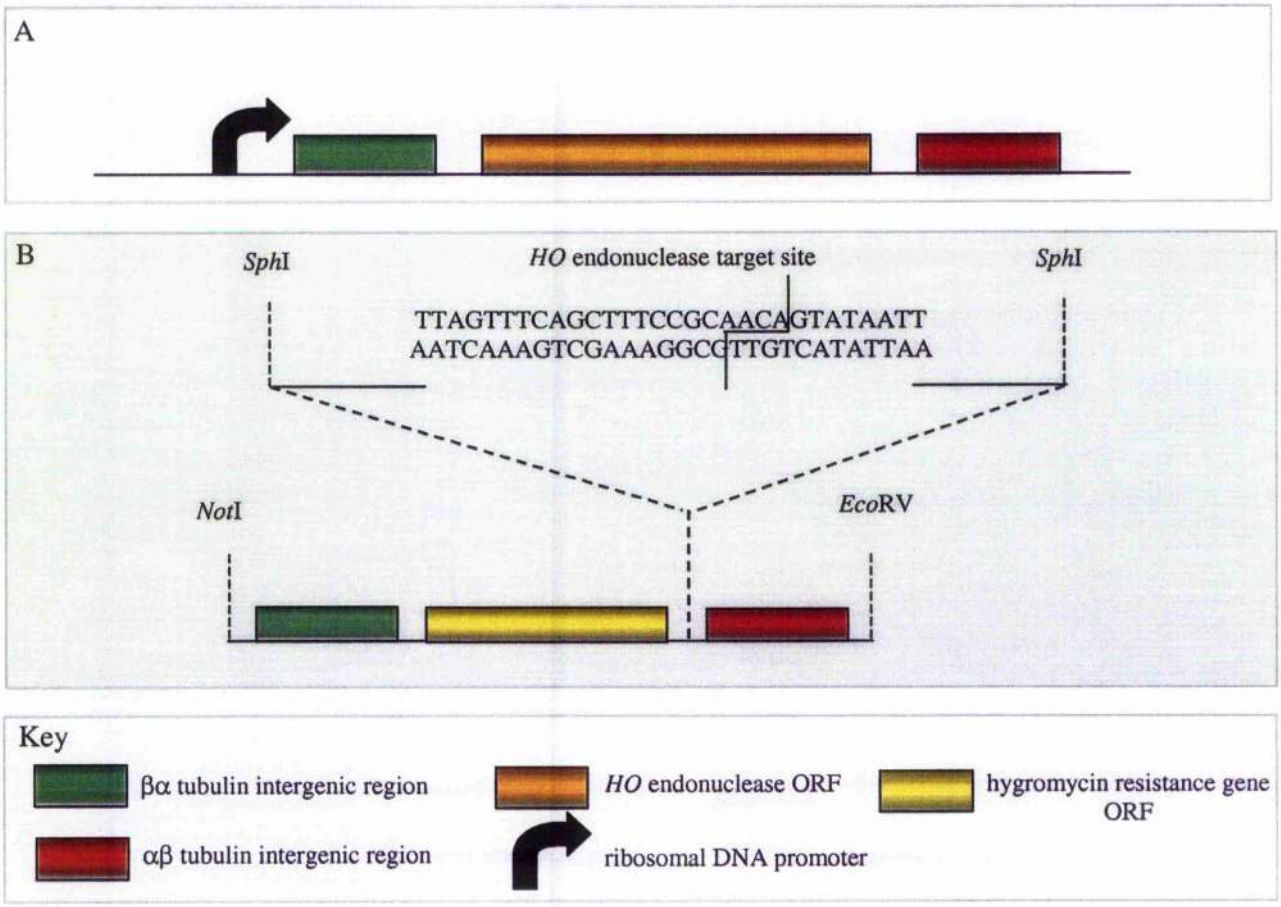

Figure 41. *HO* endonuclease experiment constructs. Expression of the *HO* endonuclease occurs from the plasmid pCC151 (A). pCC151 contains the gene ORF encoding the endonuclease flanked by tubulin intergenic processing signals. The ribosomal DNA promoter lies immediately upstream of the βα tubulin intergenic region. The HO endonuclease target site was introduced to the *T. brucei* genome using the targeting cassette pCC201 (B). This time the tubulin intergenic regions were used as targeting flanks, allowing integration into the tubulin array. The endonuclease target site was introduced into a *Sph*I restriction enzyme site immediately downstream of the hygromycin resistance gene.

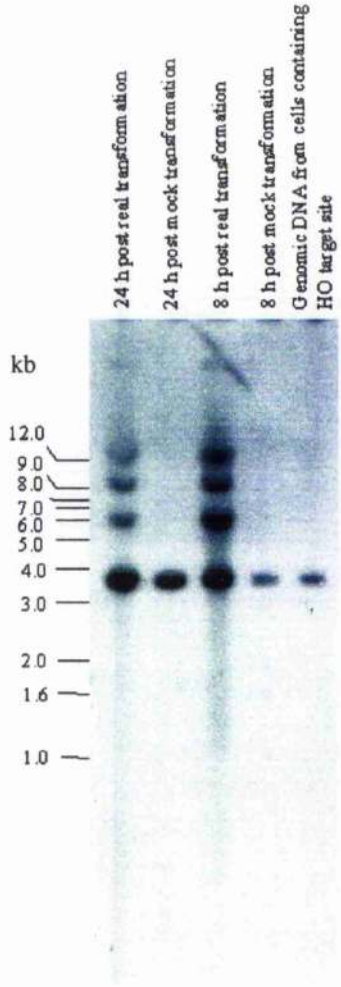

Figure 42. Transient transformation of trypanosomes containing HO target site with pCC151. Genomic DNA was prepared from cells having been transformed (real or mock) with pCC151. The preparations were digested with *Eco*RV, gel electrophoresed and Southern blotted. The blot was probed with radiolabelled hygromycin resistance gene ORF sequence and washed to a final stringency of 0.2 x SSC, 0.1% SDS at 65°C. Molecular markers are indicated to the left of the panel.

One possible explanation for no apparent DNA sequence specific digestion is that cells that have been actively transformed with pCC151 may retain the plasmid intracellularly for only short periods of time. By the time genomic DNA has been extracted, the transforming plasmid may be lost, allowing cellular recovery and DSB repair. To narrow the possibility of plasmid loss in the transformed population, genomic DNA was harvested 8 h after transformation. This was deemed an acceptable time for population recovery and at this time-point it was just feasible to extract a sufficient amount of DNA for the experiment. The experimental conditions otherwise were the same as before. Once again, only one dominant endogenous hybridising band is obvious from the Southern blot, suggesting either that repair of DSBs induced by this process is instantaneous or that the DSBs are not produced following electroporation (Fig 42).

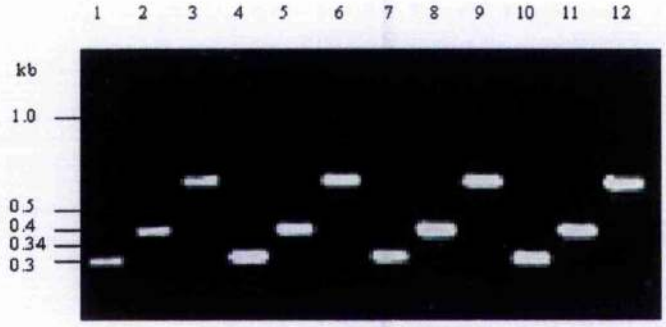

Figure 43. PCR analysis of HO target site. PCR was carried out using standard conditions. Lanes 1, 4, 7 and 10 use the HO5'P (GGCGTATATGCTCCGCATTG) and *Mat*αREV oligonucleotide combination for amplication. Lanes 2, 5, 8 and 11 use the αβ3'P (CGACGATTTCGCGCATAATT) and *Mat*αFOR combination, while lanes 3, 6, 9 and 12 use HO5'P and αβ3'P. PCR products were fractionated on a 0.7% agarose gel and stained with EtBr. Lanes 1-3 represent PCR on DNA harvested 24 h after trypanosomes were transformed with pCC151. Lanes 4-6 represent PCR on DNA harvested 24 h after trypanosomes were mock transformed with pCC151. Lanes 7-9 represent PCR on DNA harvested 8 h after trypanosomes were transformed with pCC151. Lanes 10-12 represent PCR on DNA harvested 8 h after trypanosomes were mock transformed with pCC151. Molecular markers are indicated to the left of the panel.

To ensure that the HO endonuclease target site had, indeed, been introduced to the correct locus, three independent PCR reactions were carried out with oligonucleotides

originally used to create the HO target site (*Mat*αFOR and *Mat*αREV), and oligonucleotides flanking this sequence on either side (HO5'P and αβ3'P). Three independent reactions were performed on the following four samples: genomic DNA harvested 8 or 24h after the mock transformation, and genomic DNA harvested 8 or 24h after the real transformation. The three independent reactions for each sample used the following oligonucleotide combinations: HO5'P and αβ3'P, HO5'P and *Mat*αREV, and αβ3'P and *Mat*αFOR. Products of the correct size were amplified in each case (Fig 43), confirming the locus into which the target site had been integrated and also confirming the absence of any mass insertions or deletions resulting from NHEJ repair of a DSB.

5.4 DNA repair of exogenous DNA in *rad51* mutants

RAD51 was previously disrupted in an S427 *T. brucei* strain with cassettes conferring puromycin and bleomycin resistance. This imparted a DNA repair deficiency upon cells (McCulloch & Barry, 1999). These mutants were used as electroporation targets for cassettes with non-compatible ends (i.e. non-homologous to neighbouring regions of the genome). Wild-type trypanosomes naturally integrate cassettes containing compatible ends into sequences homologous to those harboured at cassette termini. The fate of cassettes lacking terminal compatibility was not known, but provided an appropriate means of looking for other forms of DNA repair in trypanosomes, particularly in cells lacking fundamental homologous recombination machinery.

A cassette was developed using the plasmid pCP101. In order to allow activation of the selectable marker in non-transcribed regions of the genome, the ribosomal DNA promoter was cloned into an *Xba*I site, immediately upstream of the βα tubulin intergenic region. Immediately downstream of the βα tubulin intergenic region is the blasticidin resistance gene (*BSR*), actin intergenic and αβ tubulin intergenic regions respectively. The plasmid was digested with *Xho*I and *Eco*RV, releasing the pBluescript backbone along with all the components previously cloned into the multi-cloning site apart from the αβ tubulin intergenic region (Fig 44). This ensured that the cassette contained non-compatible ends, with the terminal sequences harbouring respectively pBluescript backbone and actin

intergenic DNA sequences. 5 x 10^7 trypanosomes were electroporated with approximately 5 μg of purified digested plasmid. *rad51* homozygous mutant and wild-type cells were used in this experiment. Transformants were selected in HMI-9 medium containing 10 $\mu g.ml^{-1}$ blasticidin. The process used for this selection differed slightly from previously described selection procedures. Rather than semi-solid HMI-9 agarose plates, 24-well plates were used to propagate transformants. This procedure is more conducive towards selecting slower-growing transformants (*rad51* mutants grow at a lower rate than wild-type cells).

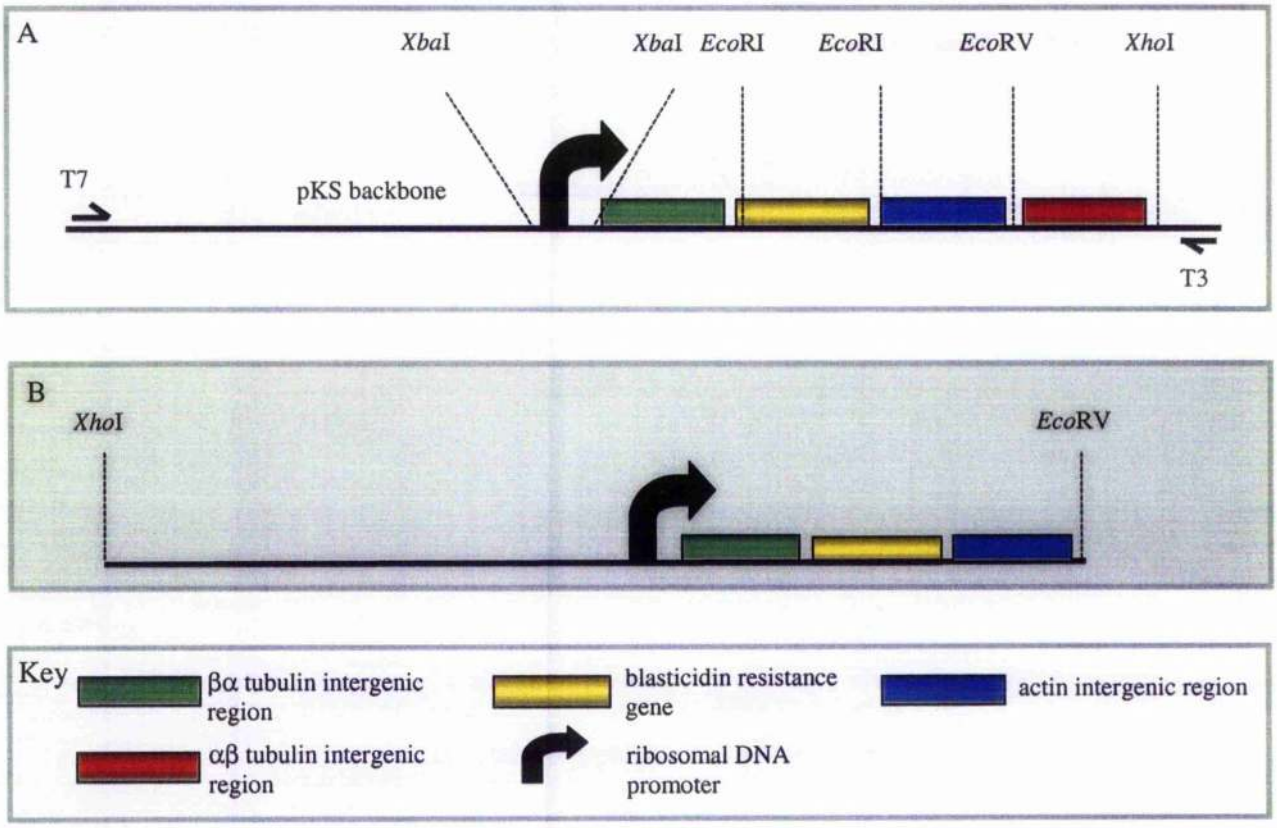

Figure 44. pCP101 and targeting cassette. Plasmid pCP101 (A) contains the blasticidin resistance gene (encoding Blasticidin S Deaminase) surrounded by actin and tubulin intergenic regions. Immediately downstream of the actin intergenic region lies the αβ tubulin intergenic region. Transcription of the blasticidin resistance gene is driven off the ribosomal DNA promoter. pCP101 contains the backbone of the bacterial replicative vector pBluescript. Restriction digestion of pCP101 with *Eco*RV and *Xho*I (B) releases the targeting cassette containing non-compatible DNA ends (i.e. pBluescript and actin intergenic region).

Following selection, the transformation efficiencies of both *RAD51* wild-type and *rad51* homozygous mutant cell lines were calculated. There was a 12-fold difference (increase) in the transformation efficiency of the null *rad51* mutants (1.2×10^{-7}) when compared with wild-type cells (0.1×10^{-7}), suggesting an up-regulation of this alternative form of DNA repair. In order to analyse the form of repair that had taken place, genomic DNA from each of the eleven mutant transformants was digested with *Sac*II, separated on a 0.7% agarose gel and the gel was then Southern blotted. The filter was hybridised overnight with α-^{32}P radiolabelled *BSR* sequence at 65°C and subsequently washed to a final stringency of 0.2 x SSC, 0.1% SDS at 65°C. Apart from clonal similarities in the hybridising restriction fragment (i.e. similar restriction fragment sizes between genomic DNA samples harvested from colonies after the same transformation), each clonal genomic DNA sample hybridised to a different size fragment (Fig 45). This, however, did not necessarily mean that each integration event had not used the same stretch of targeting DNA. Resection of targeting cassette ends may have caused variation in clonal restriction fragment size. Additionally, targeted sequence itself may have been modified, resulting in alterations to the genomic environment.

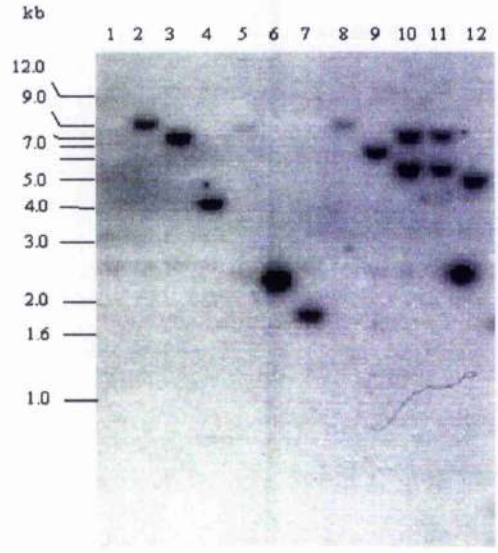

Figure 45. Genomic analysis of pCP101 transformants. Genomic DNA was extracted from trypanosome transformant clones electroporated with the plasmid pCP101. The DNA preparations were subsequently digested with *Sac*II and size-fractionated on a 0.7% agarose gel. The gel was Southern blotted and probed with α^{32}P radiolabelled *BSR* ORF sequence. This was washed to a final stringency of 0.2 x SSC, 0.1% SDS at 65°C. Lanes 1-12: 3174.2 wt, *rad51*$^{-/-}$ clones 1.1, 2.1, 3.1, 6.1, 6.2, 7.1, 8.1, 9.1, 10.1, 10.2 and 11.1. Clones 6.1 and 6.2 were derived from the same transformation as were 10.1 and 10.2. Molecular markers (kb) are indicated to the left of the panel.

In order to analyse whether integration of the pCP101 cassette produced any gross chromosomal rearrangements (GCRs) in genomes of the mutant lines, and to ensure that integrations had taken place into different loci for each independent mutant transformation, pulsed field gel analysis was undertaken. A run of 6 days was used to separate megabase chromosomes of nine of the original eleven transformants along with a *RAD51* wild-type transformant. Fig 46A shows the mega-chromosomal separation on an agarose gel. This gel was then Southern blotted using standard blotting procedures and the filter was probed with radiolabelled *BSR* DNA sequence, to a final stringency of 0.2 x SSC, 0.1% SDS at 65°C. Hybridising chromosome sizes varied between transformant clones and several hybridising bands matched aberrant chromosomes (Fig 46). This is due either to clones integrating targeting cassette sequence to different regions of the genome by illegitimate recombination reactions, or Rad51 independent reactions repairing a chromosome break by using another chromosome arm as a template for break-induced replication (Malkova *et al.*, 1996). Either form of repair by integration results in mass addition or deletion of DNA sequence to the native chromosomes that integrated the cassette.

To test whether aberrant chromosomes were formed due to cassette integration or inherent mutant inability to accurately repair DNA DSBs, *rad51* homozygous mutant and wild-type cell lines were cloned. This cloning procedure involved plating 96 trypanosomes onto a 96 well plate, theoretically resulting in clonal growth in each well. Genomic DNA plugs were then prepared from mutant and wild-type lines and chromosomes were separated as above (data not shown; work done by N.P Robinson). A lack of gross differences between mutants and wild-type clones provided evidence that the GCRs observed in the previous PFGE analysis were caused as a result of transforming *rad51* mutants with pCP101, rather than just mutating *RAD51*.

To ascertain whether the cassette could have used pBluescript sequence for a homologous recombination integration, its entire sequence was blast-searched against the Sanger *T. brucei* genomic database. Surprisingly, this gave a substantial sequence identity hit covering a long region of the pBluescript vector (93% identity over >600 bp). We believed this to be the erroneous retention of cloning plasmid sequence in the

trypanosome database and confirmation was obtained following stripping the filter from Figure 45 with 0.1% SDS and probing with pBluescript sequence (see Fig 47). The blot was washed to a final stringency of 0.2 x SSC, 0.1% SDS at 65°C and indicated there to be no pBluescript sequence in the *T. brucei* genome (lane 1). This hybridisation also revealed that a number of the transformant clones apparently had maintained some pBluescript sequence, but 6 clones had lost this sequence upon integrating the plasmid.

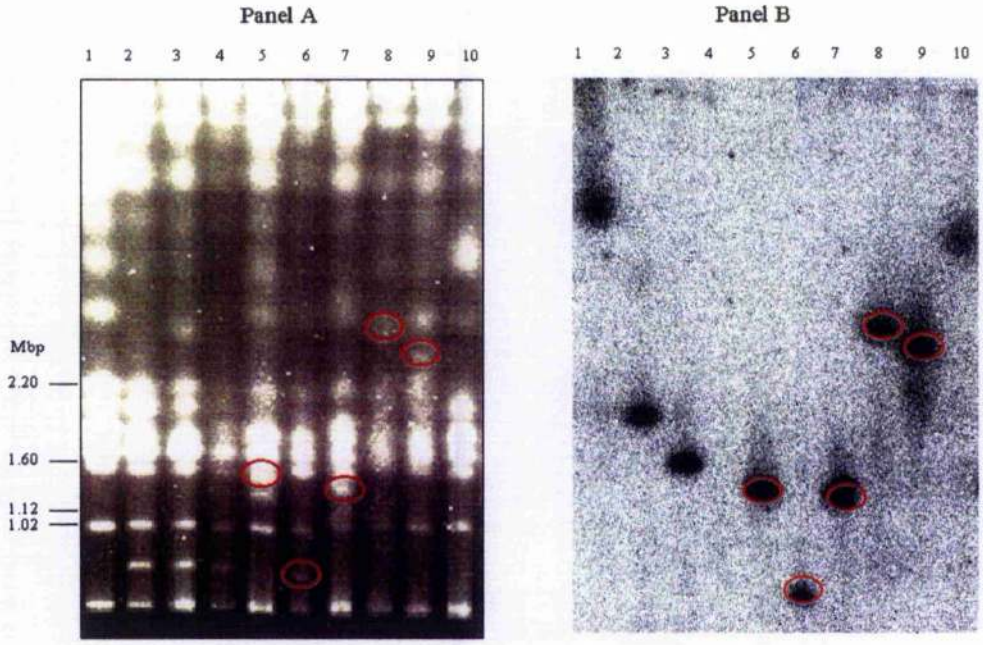

Figure 46. Pulsed field gel analysis of *rad51*$^{-/-}$ mutants transformed with pCP101. Pulsed field gel electrophoresis (6 day run) was used to separate chromosomes from *rad51* homozygous mutants transformed with linearised pCP101. The gel was Southern blotted and probed with α-P^{32} radiolabelled blasticidin resistance gene ORF DNA sequence (B). Hybridisation was carried out at 65°C and the filter was then washed to a final stringency of 0.2 x SSC, 0.1% SDS at 65°C. Several hybridising bands from the Southern blot are highlighted along with counterpart aberrant chromosome bands on the pulsed field gel (red). Lanes 1-10: wild type transformant, *rad51*$^{-/-}$ transformants 1.1, 2.1, 3.1, 6.2, 7.1, 9.1, 10.1, 10.2 and 11.1.

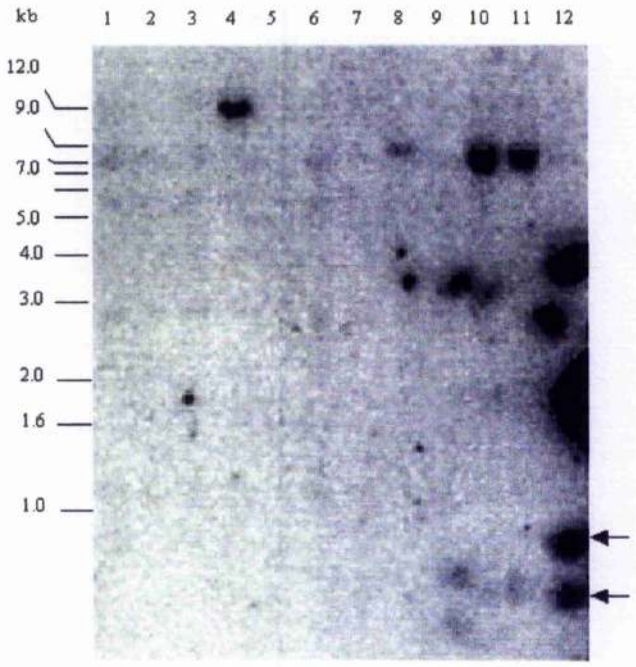

Figure 47. Presence of pBluescript in *rad51*$^{-/-}$ transformant DNA preparations. The filter from Figure 45 was stripped using 0.1% SDS and then re-probed with α^{32}P radiolabeled pKS backbone sequence. This was washed to a final stringency of 0.2 x SSC, 0.1% SDS at 65°C. Lanes 1-12: 3174.2 wt, *rad51*$^{-/-}$ clones 1.1, 2.1, 3.1, 6.1, 6.2, 7.1, 8.1, 9.1, 10.1, 10.2 and 11.1. Clones 6.1 and 6.2 were derived from the same transformation as were 10.1 and 10.2. Molecular markers are indicated to the left of the panel. Lane 12 contains spill over hybridisation products from an adjacent lane containing DNA marker which hybridises to the radiolabelled pBluescript probe. What are believed to be genomic DNA digest hybridisation products from lane 12 are arrowed to the right of the panel.

To attempt to study points of cassette integration, plasmid rescue was carried out for those 4 transformants known to contain pBluescript DNA sequence. It was hoped that several of these transformants would contain bacterial origins of replication so that exogenous DNA integration sequences following bacterial transformation with trypanosome transformant DNA preparations could be obtained. Various restriction enzymes were used separately to digest genomic DNA harvested from *rad51* homozygous mutant transformants. A self-ligation reaction was then carried out for 4 h at room temperature, the DNA was transformed into XL-1 blue MRF (supercompetent restriction

plasmid minus bacterial cells (Stratagene)) and potential transformants were selected on L-agar plates containing ampicillin. Transformants were analysed for each integration clone and in each case pBluescript by itself appeared to be the transforming plasmid, indicating contamination. A new method had to be found to recover DNA sequences encompassing points of integration in transformant cell lines. One such method, used to obtain unknown DNA sequence, is Topo walking (Invitrogen), which enables the amplification and sequencing of unknown DNA through the use of adjacent DNA of known sequence. This was originally used to obtain sequence downstream of the integrated cassette. For each of the 11 selected transformants DNA was digested with *Hae*II, dephosphorylated with Calf Intestinal Phosphatase and primer extension reactions were performed with the oligonucleotide Link-Amp I. Following oligonucleotide extension and linking reactions using topoisomerase, PCR was used with oligonucleotides specific to cassette and linker DNA (BSD FOR and Link-Amp I). Products were size-fractionated on a 0.7% agarose gel (Fig 48) and the dominant product in each lane was subsequently gel extracted (Qiagen). Products for four of the original integrants were cloned using a Topo-cloning kit (Invitrogen) and their size analysed by restriction enzyme digestion. Clones containing inserts of the correct size were sequenced and the sequence obtained was analysed to ensure it was linked to cassette DNA. Once this link had been established, sequence was blast-searched against the Sanger database using DNA:DNA alignments.

Each sequence obtained from blast searches was compared with that from the cassette terminus used in the homologous/non-homologous recombination reaction. It was found that, in each case, a stretch of homologous sequence 7 to 12 bp in length had been used to initiate an illegitimate recombination reaction (Fig 49). These reactions therefore had used what has previously been described as microhomology. It was also noticeable that every clone analysed used different regions of cassette DNA to permit integration, with one clone having removed the cassette terminus as far back as the actin intergenic region, while maintaining splicing signals.

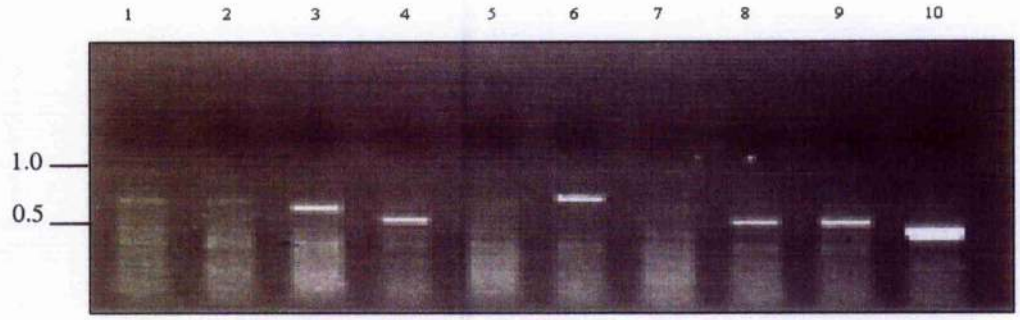

Figure 48. Analysis of DNA sequence flanking sites of pCP101 integration into *rad51* homozygous mutants. PCR products resulting from the topo-walking clones were separated on a 0.7% agarose gel and stained with EtBr. Lanes 1-10: *rad51*$^{-/-}$ clones 1.1, 2.1, 3.1, 6.1, 6.2, 7.1, 9.1, 10.1, 10.2 and 11.1. All the dominant products were between 0.5 and 1.0 kb in size. Relevant molecular markers are indicated to the left of the panel. Molecular markers are indicated to the left of the panel (kbp).

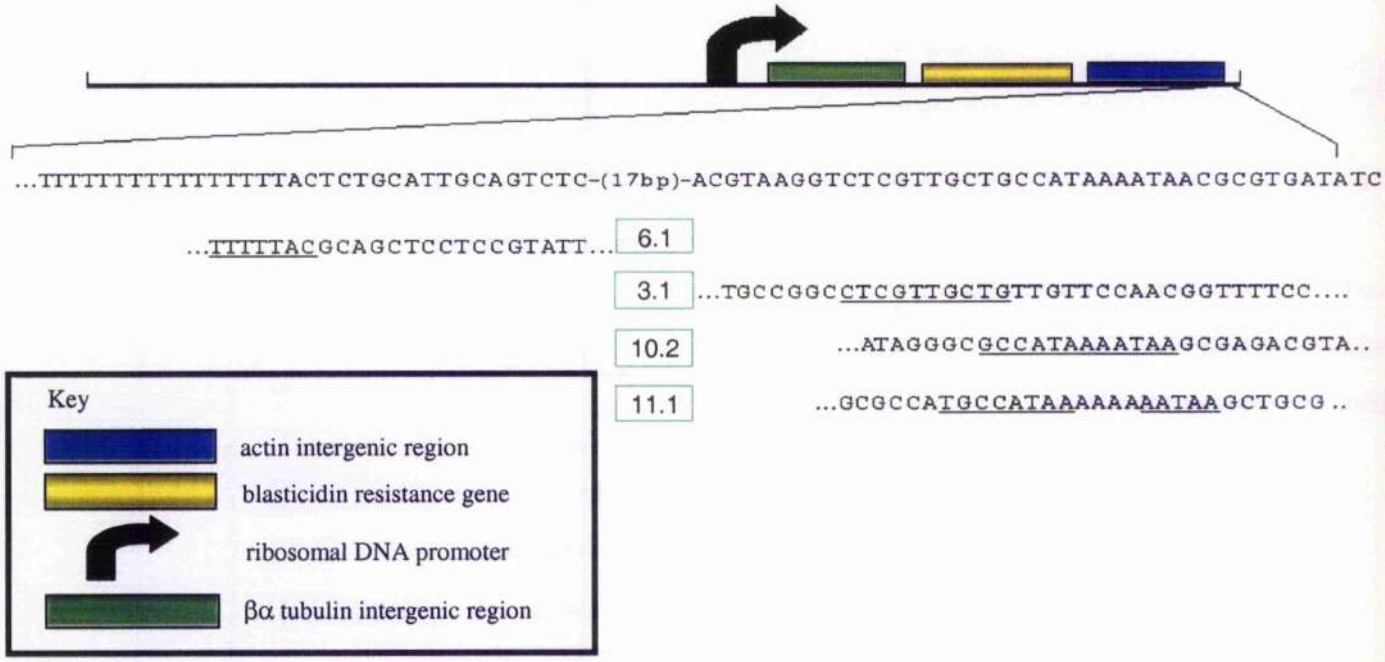

Figure 49. Aberrant integrations rely upon regions of DNA microhomology. Sequence was obtained using topo-walking (Invitrogen) for four transformant clones (*rad51*$^{-/-}$ clones 6.1, 3.1, 10.2 and 11.1) downstream of the integrated cassette, pCP101. The top line of sequence is from the integration cassette. The four other sequences show the loci into which integrations had occurred (derived from blast searches). Underlined DNA sequence represents identity between cassette and the genomic site of insertion from transformant clones.

CHAPTER 6

DISCUSSION

6. Discussion

Both Ku subunit amino acid sequences differ significantly from one organism to another (Dynan and Yoo, 1998), while polypeptide sequence lengths for both proteins are relatively consistent. Interestingly, *KU* orthologues isolated and cloned from the trypanosome genome, following database searching, revealed a reversal in the expected sequence lengths for both *KU70* and *KU80*. The trypanosome *KU70* orthologue, following double-stranded DNA sequencing, revealed a gene ORF encoding a protein with a predicted molecular weight of 81 kDa, while the putative *KU80* orthologue encodes a protein with a predicted molecular weight of 69 kDa. Although, as mentioned above, these predicted protein weights are not canonical when compared with vertebrate orthologues, lower eukaryotic Ku orthologues are less rigid in predicted weight sizes (Gell & Jackson, 1999).

On the amino acid content level, homology between the two trypanosome Ku orthologues and orthologues from both vertebrates and invertebrates is limited. This is not entirely surprising, considering the lack of immediate homology between orthologues across species from which heterodimeric partners have previously been isolated. Designation of both Ku orthologues as either Ku70 or Ku80 is due to the outcome of blastp searches, whereby the trypanosome Ku70 orthologue was placed routinely closer to other Ku70 orthologues, while the trypanosome Ku80 orthologue was placed closer to Ku80 orthologues from other organisms. Because whole protein homology alignments are not an effective guide to Ku orthologue status, regions of significant homologies have been identified across both Ku70 and Ku80 (Gell and Jackson, 1999). These regions are defined as primary homology regions (PHRs), of which there are five in each protein. Interspersed between these regions are more divergent homologous regions described as secondary homology regions (SHRs). PHRs are also reasonably conserved between both proteins, indicating that the Ku heterodimer may indeed have arisen from a homodimeric progenitor (Gell & Jackson, 1999). It proved possible to align homology regions of each putative trypanosome Ku protein with those of other orthologues using Clustal W analysis. Also, there is no C-terminal DNA-PK_{cs} interaction domain in the trypanosome Ku80 orthologue. This has been found in various vertebrates but is not evident in yeast

or *C. elegans* (Gell & Jackson, 1999). Indeed, there is no evidence for the existence of DNA-PK_{cs} in either organism. Approximately 90% of the trypanosome genome has already been sequenced and there is no significant DNA-PK_{cs} orthologue. Finally, Ku70 contains a partial putative DNA binding SAP domain (observed in other Ku70 orthologues) at its carboxy terminus (Aravind & Koonin, 2000) that has the secondary structure potential of other SAP sequences.

To analyse the function of Ku, initially looking at generally observed and previously described phenotypes and then asking whether Ku has a role in antigenic variation, we deemed it necessary to delete both alleles of either gene. The initial observation after disruption of Ku in trypanosomes was the lack of any apparent growth impediment and viability at 37°C comparable with wild-type cells. Additionally, these cells appeared morphologically identical to wild-type cells. Disruption of Ku in *S. cerevisiae* results in temperature-sensitive growth defects (Boulton & Jackson, 1996). Strains disrupted for *YKU80* are unable to form colonies on solid medium at 37°C. Mice deficient in either Ku70 or Ku80 are small, being approximately half the size of their heterozygous littermates (Nussenzweig *et al.*, 1996). These mice also display a SCID like phenotype reminiscent of defects in DNA-PK_{cs}. This is due to a deficiency in V(D)J recombination, resulting in loss of functional antigen receptors on the surface of both B- and T-cells. These defects appear to be due to dysfunctional DNA repair by NHEJ. Conversely, Ku-deficient *S. pombe* strains are perfectly viable and do not display temperature sensitivity (Baumann & Cech, 2000). The gross result of Ku deficiency in mice compared with what is observed in fission yeast and trypanosomes is likely to be due to the relative importance of NHEJ as a form of DNA repair in higher eukaryotes (see below). Indeed, NHEJ appears to be important in haploid yeast cells only in the G1 phase of the cell cycle, when no homologous chromosome exists for homologous recombination (Siede *et al.*, 1996).

In addition to examining growth of *T.brucei ku* mutants, we have determined whether or not the protein contributes to DNA repair processes. Ku is a central determinant of repair of DNA breaks by NHEJ, which is distinct from homologous recombination, the process that copies sequence from intact chromosomes. NHEJ reactions, in contrast, directly religate broken DNA ends, typically with minimal loss of intervening sequence (Lieber,

1999), depending on the species (Manolis *et al.*, 2001). The role of Ku is in recognising and binding the broken DNA ends, with at least three other conserved factors needed to catalyse break repair in higher eukaryotes: DNA-PK_{cs}, Ligase IV and XRCC4 (Featherstone & Jackson, 1999). As mentioned above, DNA-PK_{cs} has not been identified in the complete genome sequences of either *S. cerevisiae* or *S. pombe* (Manolis *et al.*, 2001).

We found no difference between wild-type, heterozygous or homozygous *ku* mutants in their sensitivity to the radiomimetic compound methyl methanesulphonate, an alkylating agent that causes single- and double-strand DNA breaks. This is in contrast to *rad51* mutants, which display increased MMS sensitivity in the same range of drug concentrations (McCulloch and Barry, 1999). It appears, therefore, that if Ku plays a role in repair of trypanosome DNA damage it is masked by the greater prevalence of homologous recombination. This bears comparison with yeast *ku* mutants, where no increase in DNA damage sensitivity is detectable (Boulton and Jackson, 1996; Mages *et al*, 1996; Siede *et al*, 1996; Boulton & Jackson, 1998; Manolis *et al*, 2001), and the background role played by Ku in *S. cerevisiae* DNA repair is seen only when the homologous recombination gene *RAD52* is also inactivated.

One reason for the apparent absence of DNA repair defects in trypanosome *ku* mutants is the assay employed to analyse DNA DSB repair. Higher eukaryote cell lines deficient in DNA-PK_{cs} or Ku80 have defective DNA DSB rejoining and are extremely sensitive to ionising radiation and other agents that directly generate DNA DSBs. In contrast, they are much less sensitive to agents such as ultraviolet light, alkylating agents and mitomycin C that produce other forms of DNA damage (Collins *et al.*, 1993). As stated above, the agent used to test for DNA repair disabilities in trypanosome mutants was methane methyl sulfonate, which is an alkylating agent.

Various assays have been developed to look for the presence of DNA repair by NHEJ (see below), but none has revealed the presence of this process in trypanosomes. It remains possible, however, that Ku is involved in NHEJ in trypanosomes. For instance, it is possible that DNA integration or nuclear entry following transformation occurs only at certain phases in the cell cycle, where DNA replication and cell division occur, and where homologous exchange in higher eukaryotes is thought to assume importance.

Perhaps trypanosome NHEJ is limited to G1 phase, when no appropriately positioned sister chromatids are available as substrates for homologous recombination. Alternatively, NHEJ may be limited to non-dividing trypanosome life cycle stages, when limited DNA replication occurs. These stages are used in trypanosomes for transmission from mammals to the tsetse fly and vice versa (see Barry and McCulloch, 2001 for review). It is interesting in this regard to note that murine *LIGIV* and *XRCC4* mutants display neuronal cell death during gestation (Frank *et al*, 1998; Gao *et al*, 1998).

A clear phenotype in the *ku* homozygous mutants is defective telomere length maintenance. The extent of telomere attrition in the absence of Ku is more pronounced in *T. brucei* than in the same mutants in yeasts (Baumann & Cech, 2000; Boulton & Jackson 1996). Telomeric sequence loss in *S. pombe ku* mutants results in telomeres that are ~100 bp shorter than wild-type counterparts, while *S. cerevisiae ku* mutants differ from wild-type cells by approximately 300 bp (Baumann & Cech, 2000; Boulton & Jackson 1996). A similar telomere shortening phenotype is observed when *RAD50*, *MRE11* and *XRS2* are mutated in *S. cerevisiae* (Boulton & Jackson, 1998). The same effect is not imparted on *S. pombe* upon mutation of *MRE11* (Manolis *et al.*, 2001). Trypanosome telomeric tracts, however, are substantially longer than yeast sequences and exhibit a higher level of heterogeneity. In mammalian cells, inactivation of Ku has a less pronounced and opposite effect to that seen here, with moderate lengthening (Samper *et al.*, 2000). It has been proposed that the observed difference may result from the absence of an active DNA-PK_{cs} orthologue in yeast (Boulton & Jackson; 1996).

Quantification and comparison of the extent of telomere loss in *ku* null mutants and wild-type and heterozygous clones proved difficult, due to the large terminal restriction fragment sizes involved. Re-introduction of *KU70* to its original locus in *ku70* homozygous mutants allowed us to assess accurately telomere expansion starting from a lower average terminal restriction fragment (TRF) basal level, giving a clearer picture of the extent of this phenotype. TRFs from transcriptionally active telomeric loci in null mutants were, on average, 2.9kb shorter than those observed in *KU70* re-expressors. Interestingly, this size variation was substantially lower than that of TRFs from transcriptionally inactive telomeric regions; following re-introduction of *KU70* they had an average increase of only 400 bp. In wild-type trypanosomes, transcriptionally active

telomeres have previously been demonstrated to grow at a faster rate than inactive telomeres (Pays *et al.*, 1983c). Re-expressors harbouring particularly short telomeres (~ 1 kb) extend at a prolific rate (~ 170 bp / generation) when compared with re-expressors starting with long stretches of telomeric DNA (> 7 kb). The rapid expansion of short transcriptionally active telomeres is analogous to rapid telomere generation observations made from newly formed telomeres in *T. brucei* (Horn *et al.*, 2000).

No single clone analysed appeared to have lost all telomeric sequence. It is noticeable that null mutant clones, exhibiting TRFs as small as 4 kb (of which the telomere tract comprises ~2 kb), are viable and grow at the same rate as wild-type cells, demonstrating that loss of telomeric sequence does not inevitably lead to decrease in cell viability. Similarly, *ku* mutants in *S. pombe* maintain viability comparable with that of wild-type cells (Manolis *et al.*, 2001). However, disruption of Ku in budding yeast results in telomere shortening and temperature sensitivity of growth (Boulton & Jackson 1996). Retention of normal cellular function, in spite of telomere attrition, poses the quandary of the requirement for trypanosomes to maintain relatively extensive telomere length, at least in the active BES. The choice strain in this analysis, S427, is a laboratory-adapted strain amenable to a variety of experiments. It may be the case that, in the field, faster switching parasites require more extensive telomeric tracts as substrates for catalysing *VSG* switching recombination reactions; this can be tested by deleting *KU* in pleomorphic trypanosome strains.

Mice deficient in telomerase RNA ($mTR^{-/-}$) are initially viable but, after several generations, telomere shortening and chromosome fusions result in infertility and hematopoietic system defects (Blasco *et al.*, 1997). Trypanosome *ku* mutants were maintained in culture for over 150 generations, with no apparent loss in cell viability. These same mutant lines did not exhibit further loss of telomere sequence after 150 generations, with an apparent equilibrium being reached. Disruption of Ku in yeast results in earlier lethality for telomerase mutants, suggesting that, in the presence of Ku, telomeres can be maintained and stabilised (see Kass-Eisler & Greider, 2000 for review on telomere-length maintenance). Therefore, trypanosome *ku* mutants are likely to be using a second pathway for equilibration of short telomeric substrates.

Shortening of telomeric sequences through loss of Ku also suggests that chromatin packaging has in some way been compromised, allowing accessibility of telomeres to nucleases. This may simply be due to loss of nuclease blockage caused by loss of trans-acting protein assemblies involved in telomere protection, or alternatively it may be due a reduction in DNA stability through loss of DNA-DNA interactions mediated by Ku. One form of DNA complex interaction is the invasion of the 3' telomeric overhang forming a terminal loop structure known as the telomere loop (t-loop). These structures have been reported in African trypanosomes, humans and ciliates (Munoz-Jordan *et al.*, 2001; Griffith *et al.*, 1999; Murti *et al.*, 1999). In yeast, it has been proposed that telomere tracts fold back and interact with adjacent duplex DNA via protein-protein interactions (Ray *et al.*, 1999). Ku previously has been demonstrated, by crosslinking and coimmunoprecipitation experiments, to interact with the repeat-binding proteins TRF1 and TRF2, which are known to be involved in the formation of t-loops (Song *et al.*, 2000). Proposed functions of the t-loop are the protection of telomeres from nuclease attack and prevention of these ends from recognition by endogenous machineries, as DNA double strand breaks. If Ku is involved in t-loop assembly, its loss may provide a naked DNA substrate for enzymatic cleavage resulting in shortened telomeres.

The major reason for studying the function of Ku in trypanosomes is the ability of the parasite to evade host immune systems using antigenic variation. This evasion capacity is active twice in the life cycle of the parasite, in the bloodstream and metacyclic stages. *VSG* switching in bloodstream trypanosomes occurs mainly in two ways: recombinational and transcriptional. A third proposed means of effecting *VSG* switching, multiple point mutations that arise during the generation of a *VSG* gene copy (Donelson, 1995), is more contentious (Graham & Barry, 1996). As this mechanism is unlikely to form a major pathway of VSG switching, this discussion will concentrate on the roles played by transcriptional and recombinational *VSG* gene switching.

Transcriptional switching results in activation of a silent expression site in conjunction with inactivation of the previously active expression site. It has recently been shown, however, that the transcriptional activity of expression sites is not exclusive, with transcription extending for short stretches of DNA into what were previously described to be inactive expression sites (Vanhamme *et al.*, 2000). It must also be stressed that,

although tracts of "silent" expression sites do appear to be transcriptionally active, transcription does not extend as far as *VSG* gene sequences. This work suggests and somewhat agrees with previous work suggesting that all expression sites are on "stand-by" to become completely active upon the occurrence of an expression site switch (Navarro *et al.*, 1999).

The forms of regulation that are proposed to give rise to the complete transcriptional activation of a single expression site are numerous. One model that has garnered support over the years is that of telomeric silencing. This was first described in *S. cerevisiae* (Gottschling *et al.*, 1990) as the silencing effect registering from sub-telomeric loci and is believed to be due to chromatin conformations enveloping the telomere. In order to explain a potential role for Ku in antigenic variation due to telomeric silencing, it is necessary to explain both metacyclic and bloodstream expression sites (see introduction for more detailed explanation). Characterised bloodstream expression sites are 40-60 kb in size and polycistronic. Metacyclic counterparts are much smaller, consisting of monocistronic cassettes composed of up to 5 kb of sequence. Both expression cassettes are believed to be transcribed by RNA polymerase I (Chung *et al.*, 1992; Lee and Van der Ploeg, 1997) and contain either large stretches (bloodstream expression sites) or 0-2 units (metacyclic expression sites) of 70-bp repeats (Lenardo *et al.*, 1984; Liu *et al.*, 1983; Campbell *et al.*, 1984; Shah *et al.*, 1987). Previous studies of the extent of telomeric silencing effects in budding yeast has shown that DNA sequences up to 5 kb from telomeric repeats are subjects of this form of transcriptional repression (Borst *et al.*, 1998). The proximity of this effect suggests it as a possible mechanism in the regulation of *MVSG* gene promoters and ultimately *MVSG* differential activation.

The bloodstream expression site is not a strong candidate for telomeric silencing. Promoters from characterised bloodstream form expression sites lie 40-60 kb upstream of telomeric repeat tracts. This distance is probably too great to be affected by classical telomeric silencing. Recent work has argued against the simple model of telomeric silencing in trypanosomes. There is no obvious difference in chromatin structure around the promoter region between active and inactive expression sites (Navarro & Cross, 1996). Also, manipulation of an inactive expression site, such that a bacteriophage T7 RNA polymerase promoter was substituted for the endogenous expression site promoter,

rendered the expression site transcriptionally active (Navarro *et al.,* 1999). The same paper points out that this expression site is inactive upon differentiation to the procyclic stage, suggesting an accessible chromatin structure at inactive expression sites in the bloodstream stage of the parasite. More recently, it has been demonstrated that the active expression site is localised to an exclusive nuclear transcriptionally active domain (Navarro & Gull, 2001). This suggests that, rather than a system of locus -specific silencing, a mono-allelic transcriptionally poised body compartmentalises transcription machinery to the prospectively active expression site. To test further for the occurrence of telomeric silencing, components of the Ku heterodimer are an obvious choice for investigation, with the trypanosome being a good model, harbouring native genes sub-telomerically.

Deletion of Ku from the trypanosome appears not to disrupt regulation of the transcriptional status of genes in sub-telomeric loci. In budding yeast, Ku is localised primarily at telomeres (Gravel *et al.,* 1998) and Ku-deficient strains lose telomeric silencing (LaRoche *et al.,* 1998). Ku-dependent TPE is not, however, observed in *S. pombe* (Manolis *et al.,* 2001) leading to the conclusion that both processes are not always intrinsically linked. As with trypanosome Ku, when members of the Mre11 complex are mutated in *S. cerevisiae* no derepression of silencing occurs in conjunction with telomere shortening (Boutlon & Jackson, 1998).

Analysis of *MVSG* expression sites showed no alleviation of *MVSG* silencing. This experiment, however, is limited by the life cycle stage at which differential *MVSG* expression can be assessed. We looked for the effect on activity in expression sites in the procyclic stage of the trypanosome life cycle, although *MVSG* genes are active in the metacyclic and early bloodstream stages. This lack of repression of silencing in *ku* mutants indicates that *T. brucei* Ku is functionally more related to its fission yeast orthologue than to the *S. cerevisiae* orthologue although other factors may cause this difference in function. As yet, there has been no formal demonstration of direct protein complex involvement in silencing transcriptionally inactive expression sites in trypanosomes, though recent work by Navarro & Gull (2001) indirectly points to such a complex. It may indeed be the case that a novel network of proteins are involved in cross-talk and regulation of antigenic variation in *T. brucei.*

We find no evidence that Ku plays a role in either recombination or transcriptional *VSG* switching, since the mutants display wild-type levels of switching and a similar profile of switching reactions. This is in keeping with the hypothesis that recombinational *VSG* switching involves homologous recombination, rather than other forms of DNA exchange (McCulloch and Barry, 1999).

Telomeres are integral in nuclear architectural maintenance and are observed as clusters at the nuclear periphery. This type of localisation is believed to play an important role in DNA recombination and telomeric silencing (Gotta *et al.*, 1996; Cockell *et al.*, 1999). Indeed, TPE is associated with chromatin organisation, whereby proteins involved in telomeric silencing are localised to sub-telomeric and telomeric regions. Recent studies have focused on the interactions between telomere protein complexes and the nuclear periphery. Galy *et al.*, (2000) studied nuclear pore complexes (NPC) and in particular two proteins involved in complex formation, Mlp1 and Mlp2. These complexes are involved in RNA and protein transport between the cytoplasm and the nucleus. Mutation of these two proteins in *S. cerevisiae* does not result in disabled nuclear transport (Galy *et al.*, 2000). However, a phleomycin sensitivity phenotype associated with disruption of DNA repair was observed. It was subsequently demonstrated that Mlp1 and Mlp2 are indeed involved in repair of DNA double-stranded breaks (Galy *et al.*, 2000). This surprising observation led Galy *et al.*, (2000) to look at telomere binding proteins that are concurrently involved in DNA repair. Obvious candidates are Ku70 and Ku80. Immunofluorescence showed that the localisation of Ku, upon disruption of Mlp2, is dramatically altered. Ku is normally localised to telomeric loci, but Mlp2 disruption results in distribution of Ku throughout the nucleus. Indeed, disruption of Mlp2 also causes the repression of telomeric silencing, intrinsically linking nuclear localisation with transcriptional status (Galy *et al.*, 2000).

With the transcriptional status of sub-telomeric regions harbouring *VSG* genes being our priority, we decided to investigate the role if any, Ku has, in the localisation of telomeres within the trypanosome nucleus. Our observations that Ku70 is not important in telomeric silencing of genes naturally occurring in sub-telomeric regions of the trypanosome genome led us to the intriguing question whether trypanosome Ku might influence nuclear localisation yet have no apparent role in TPE and *VSG* gene activation.

We addressed this question using fluorescent in situ hybridisation (FISH) using both telomere and mini-chromosomal specific probes. There was, however, no deviation in the localisation of telomeric or minichromosomal sequences upon disruption of Ku, suggesting no role in nuclear organisation. It is also worth noting that this effect was not altered by life-cycle stage, with localisation at perinuclear regions being maintained in both bloodstream and procyclic stages of the cycle.

As mentioned above, Ku is involved in the docking of telomeres to the nuclear periphery in budding yeast. The importance of this tethering reaction has not yet been realised, with several theories being proposed, including the occurrence of telomeric silencing. In trypanosomes, the telomeres of minichromosomes reside naturally at the periphery of the nucleus (Chung *et al.*, 1990). Minchromosomes act, essentially, as a reservoir for *VSG* genes. Although many *VSG*s are in minichromosomes (of which there are up to 100 in each trypanosome cell), transcription of *VSG*s can be achieved only following reactions that result in movement of *VSG*s to bloodstream expression sites. There appears to be no second firm role for minichromosomes throughout the life cycle of the cell and it has been proposed that they may in fact exist as breakdown products of intermediate- or mega- chromosomes (personal communication, G. Rudenko).

Loss of Ku from the cell could lead to mislocalisation, or ultimately absence, of these minichromosomes in trypanosome mutants. In order to address this, minichromosomes from both *ku70* and *ku80* null and heterozygous mutants were fractionated on pulsed field gels. There was no gross difference in EtBr staining intensities between mutant and wild-type lanes. The gels were subsequently Southern blotted and, when probed with a marker known to be present in minichromosomes, the filters did not display absence of marker sequence in the mutant lanes. Amalgamation of these data with the FISH localisation analysis suggests that loss of Ku does not lead to loss or, indeed, mislocalisation of minichromosomes. No effect on megachromosomal localisation was observed. This reinforced the idea that trypanosome Ku does not act as a docking protein and hence differs from its budding yeast orthologue. Similar experiments have not been carried out in fission yeast, hence preventing a comparison.

Disruption of Ku in *S. pombe* results in a further telomeric effect that essentially causes chromosomal fusions (Baumann & Cech., 2000). Quantitative FISH analysis carried out

on metaphase nuclei from primary wild-type and Ku86$^{-/-}$ mouse embryonic fibroblasts (MEFs) derived from heterozygous crosses (Samper *et al.*, 2000), revealed that Ku mutant cells had a 24-fold increase in telomeric fusions when compared with Ku wild-type cells. This phenotype far outnumbers the frequency of other chromosomal aberrations. All telomeric fusions contained, at the fusion point, telomeres of an average length of 81.6 +/- 5.3 kb, suggesting that they did not occur due to loss of telomeric sequences. Pulsed field gel separations of mega-chromosomes from trypanosome *KU* wild-type and *ku* homozygous mutant cell lines suggested that Ku-deficient trypanosomes do not contain fused chromosomes (data from N.P.Robinson). This poses the question whether trypanosome Ku is actually involved in capping chromosomal termini or in fact localises upstream to sub-telomeric regions. Absence of observed fusion reactions suggests that trypanosome Ku does not act as a terminal blockade against joining reactions or, no such end-joining reactions occur in trypanosomes.

To test whether trypanosome Ku is a DNA repair protein, we have attempted to develop a number of assays to look for DNA repair by NHEJ. One way of attempting to monitor the occurrence of DNA DSB repair by NHEJ was by introducing a previously absent restriction enzyme target site to the *T. brucei* genome. By transiently transforming trypanosomes harbouring this unique HO target site with a plasmid allowing expression of HO endonuclease, it might be possible to assess how breaks in the cells' genomes are processed. HO endonuclease initiates mating type switching, an intrachromosomal recombination event, by cleavage of DNA at a single chromosomal locus *in vivo* (Haber, 1998). Previous demonstrations have indicated a role for several proteins, including members of Rad50/Mre11, Ku and Sir complexes (Lee *et al.*, 1999; Bennet *et al.*, 1996), in repairing DSBs that are created by HO endonuclease activity.

Our attempt at using HO endonuclease as a tool introduced one possible problem. The HO endonuclease target site sequence introduced to the trypanosome is not naturally present in the *T. brucei* genome and may induce DNA structural variations to the locus of a constitutively active gene array, possibly precipitating further complexities. Another potential complexity relates to the use of transient transformation as a method of enabling expression of the endonuclease. It is, however, known that certain reporter genes can indeed be expressed transiently from a plasmid in *T. brucei* (Zomerdijk *et al.*, 1991).

Also, even upon successful expression of the endonuclease, the nature of trypanosome transformations means that only a small percentage of the population is actually transformed successfully with the plasmid. With this in mind, procyclic form trypanosomes were used as the life cycle stage of choice as their transformation efficiency is far greater than that of bloodstream form cells (10^2-10^3 fold greater transformation efficiency). Ultimately, using two different methods for break detection, there was no evidence for the occurrence of any DSBs in cell populations transformed with the HO cassette plasmid. The methods of detection used included PCR amplification and Southern blot analysis. The absence of extra bands in Southern blot analysis in transformant cell lines is not entirely surprising, considering that only a small percentage of the cells would actually express the endonuclease. Using PCR as a method of DSB repair detection would allow us to ascertain whether DNA sequence insertion or deletion arose during repair. There was no clear deviation from expected PCR products using oligonucleotides surrounding the endonuclease target site. This suggests that, if breaks are occurring, they are being processed with integrity, not grossly changing the genomic environment of the initial break.

Adaptation of another assay used in other systems to monitor DSB repair allowed us once again to look for NHEJ in *T. brucei*. This plasmid re-joining assay depends on endogenous repair machineries to effect re-ligation of exogenous linearised plasmid DNA upon introduction to the cell. As before, the assay depended on the ability of trypanosome cells to metabolise non-native DNA introduced to the cell via transformation. This experiment has intricate problems, largely due to the fact that *T. brucei* does not naturally replicate plasmid DNA. Secondly, the necessary elimination of all traces of supercoiled DNA prior to electroporation is an extremely difficult task. Contamination with this form of DNA would give rise to false positive results. However, if NHEJ were a prevalent form of DNA repair in trypanosomes, this assay would be a relatively simple way to study its efficacy. Although elimination of supercoiled plasmid DNA was almost wholly achieved in our experiment, the effects of electroporating linear DNA into trypanosomes were minimal (i.e. there was little evidence of plasmid repair), suggesting that NHEJ is either a rarely utilised form of DNA DSB repair or trypanosomes are unable to repair foreign DNA introduced to populations by electroporation. We do

know that this form of repair is used to repair plasmids introduced to yeast via electroporation (Boulton and Jackson, 1996), but that organism is fully capable of propagating plasmid DNA, and so repair of this form of DNA would be expected.

The final method utilised to look for the presence of DNA repair in trypanosomes *via* NHEJ was carried out in a homologous recombination deficient strain. It was hypothesized that such a deficiency would facilitate detection of compensatory repair pathways such as NHEJ. To test for alternative forms of DNA repair, cassettes containing non-compatible termini were introduced to the trypanosome via electroporation. The first striking observation made in trypanosome *rad51* homozygous mutants was the dramatic changes in chromosomes into which integration had occurred. This observation, made by pulsed field gel electrophorsesis, revealed newly sized chromosomes. On further Southern analysis, using as probe the transforming cassette, it became clear that chromosomes of novel size had arisen in most of the transformed clones.

Gross chromosomal rearrangements, including translocations, deletions and inversions, are commonly observed in cancer cells (Lengauer *et al.*, 1998; Padilla-Nash, 2001). Pathways involved in GCR suppression are not, however, well characterised. Generation of GCRs in cell populations generally result from mutation in DNA repair genes and genes encoding S-phase checkpoint functions in *S. cerevisiae* (Chen and Kolodner, 1999; Myung *et al.*, 2001). With this in mind, it is not surprising that the first evidence of GCRs in *T. brucei* result from disruption of the DNA repair gene, *RAD51*. It was not possible to test directly the types of arrangement that gave rise to aberrant chromosome sizes in *rad51* homozygous mutants.

Mutation of Rad51 in *S. cerevisiae* causes an increase in GCR rate (Myung *et al.*, 2001), but to a lesser extent than observed in *rad52* mutant cells (Myung *et al.*, 2001). Again, it was not directly possible to compare organisms (yeast and trypanosome), with respect to the hierarchy of proteins involved in maintaining genomic integrity. Recent work (manuscript in preparation, Robinson *et al.*,), however, has allowed some degree of comparison, with Mre11, an integral member of the homologous recombination machinery. As with *S. cerevisiae*, disruption of *MRE11* in *T. brucei* has a significantly greater effect on chromosome maintenance than disruption of *RAD51*. This result is

consistent with the effects of mutations in these genes on break-induced replication (Signon *et al.*, 2001). Also consistent with observations in *S. cerevisiae*, deletion of *KU70* appears not to have an obvious effect on genome integrity. The role of Mre11 in acting as a "guardian" of the yeast genome may be linked intrinsically to its role in break induced replication. *rad51* mutants are still capable of performing DNA repair by break-induced replication, which is believed to be due to Mre11. Participation in these different pathways could explain the difference in efficacy between Rad51 and Mre11 in maintaining the integrity of the yeast genome.

As mentioned earlier, Ku appears not to have a major role in suppressing GCRs in *S. cerevisiae* (Chen and Kolodner, 1999). This, however, does not appear to be the case in higher eukaryotic cells. *RAD54*$^{-/-}$*KU70*$^{-/-}$ double null mutants are profoundly more sensitive to spontaneous chromosomal aberrations than *RAD54*$^{-/-}$ B-cell line DT40 mutants (Takata *et al.*, 1998). This, once again, suggests more similarity in the function of Ku in trypanosomes to that of Ku in *S. cerevisiae* than in mammals.

Sequence walking was then used to deduce the nature of the regions adjoining the newly integrated cassette to attempt to interpret whether integration had occurred via stringent homologous recombination or NHEJ reactions. Initial analyses revealed, however, that small sequences of homology were used to integrate the 3' flanks of exogenous DNA cassettes. Out of the four clones analysed for exogenous DNA sequence maintained upon integration into the trypanosome genome, it became clear that in each case some form of microhomologous recombination event had taken place. The annealing of overlapping microhomologies has been suggested (Kramer *et al.*, 1994; Nicolas *et al.*, 1995) to occur similarly to the repair of DSBs by a homologous recombination mechanism known as single-strand annealing (SSA) (Fishman-Lobell *et al.*, 1992). Indeed, there are believed to be three distinct pathways involved in the suppression of gross chromosomal rearrangements in budding yeast: two that suppress microhomology (*RFA1* and *RAD27*) and one that suppresses non-homology-mediated GCRs (*RAD50/MRE11/XRS2*) (Chen and Kolodner, 1999).

Detection of extensive GCRs in trypanosome *mre11* homozygous mutants, resulting in loss and altered locus positioning of several genomic DNA markers, suggests that this protein is a likely candidate for maintenance of the trypanosome genome. One finding

upon analysis of regions adjoining integration sites in *rad51* mutants is the extent of homologous sequence required to catalyse integration reactions. We observed integrations to be associated with tracts of up to 13 bp of sequence identical to the exogenous DNA used for transformations. However, gene-targeting experiments in yeast can use homologous DNA stretches as small as 35 bp in length (Wach *et al.*, 1994). Whether this can be categorised as a classical homologous recombination reaction rather then non-homologous end joining remains to be seen. Independently, we have shown that trypanosomes can use as little as 24 bp of DNA sequence identity at either end of a targeting cassette to integrate into the genome (Conway *et al.*, manuscript in preparation). Similar homologous-nonhomologous reactions are well characterised in mammalian transformations (Haber, 2000). These reactions occur at a significantly higher rate in trypanosomes containing disrupted *rad51* than in wild-type trypanosomes.

The importance of microhomologous DNA recombination events in trypanosomes is not known, though it could be postulated that it may play some role in antigenic variation. *VSG* genes are normally arranged in tandem arrays scattered throughout the genome. Gene ORFs are preceded by a number of 70-bp repeats that act as substrates for homologous recombination and gene conversion events involved in *VSG* switching reactions. It is often the case that limited numbers of repeats are utilised in these reactions, particularly in reactions that catalyse *MVSG* switching. These repeats are often degenerate and it is not unreasonable to hypothesize that microhomologies may be important in dictating switching events.

Future work:

The hypothesis underlying these experiments was that the Ku heterodimer is involved centrally in *VSG* switching reactions either directly or indirectly. The trypanosome is potentially a good model for Ku functioning in TPE, as it contains endogenous, sub-telomeric *VSG* genes. The results to date suggest that Ku does not participate in *VSG* switching, but confirmation is necessary and ultimately switching analysis must be carried out in mutants derived from faster switching trypanosomes. This would assist in tackling the question of locus specific effects dictating VSG switching. To rule out TPE entirely, other proteins involved in eliciting this transcriptional repression must be disrupted and the effect on sub-telomeric transcriptional status monitored. Several methods may be used to isolate such factors, either using anti-Ku antibodies to precipitate further telomere binding proteins or searching the trypanosome genome database with telomere silencing protein sequences.

To analyse further whether trypanosomes are capable of performing DNA repair by conventional non-homologous end-joining reactions, new assays must be developed. One option is the utilisation of an *in vitro* assay for this form of repair involving cell-free extracts. This method involves incubation of extract preparations with linear DNA substrates. Functional proteins involved in performing NHEJ reactions subsequently catalyse rejoining events causing concatamerisation of the plasmid substrate. Although this approach is crude and not entirely indicative of endogenous processes, it would be a simple way to test whether these machineries exist.

To follow up the finding of a Rad51 independent mechanism for DNA repair in trypanosomes, candidates for catalysis must be identified and characterised. A prime candidate is the Mre11 complex. As mentioned above, DNA repair by break induced replication can be carried out in the absence of Rad51. This reaction is performed by the Mre11 complex in budding yeast (Chen and Kolodner, 1999). The ability of Mre11 to perform reactions that result in microhomologous reactions in trypanosomes could be tested in the same manner used to look for alternative forms of DNA repair in *rad51* null mutants. Insights into the importance of microhomologies in reactions that culminate in a *VSG* switch may well be gained following isolation of factors that are involved in this

form of DNA repair. As with the Ku heterodimer, any critical analysis involving factors potentially involved in *VSG* switching must be carried out in pleomorphic parasites that utilise antigenic variation more frequently than laboratory adapted strains.

APPENDICES

Appendix 1 Amino acid and DNA sequence for Ku70/*KU70*

DNA sequence was obtained following double pass sequencing of ILTat1.2 genomic lambda clones containing the *KU70* open reading frame. Sequencing oligonucleotides are in bold for KU70R1-R6 and the reverse complementary sequences for KU70T1-T7 are in red (see section 2.5.3). T3 and T7 oligonucleotides were also used for sequencing. Start codons are represented by black bold, italicised, underlined text and stop codons by red bold, italicised, underlined text.

Ku70 amino acid sequence

MAHYDEWLMSVDGLDFHDEDDSYFNETYATEDEPSLRTGSVSSL
DQFDVVLCLVDFQQRMFQGVVAHDCKNPFAGGAEQVYDDITEVKVKETPTVFEKAICC
VQQLYKDKGISDSNDMVALVLYNTRECTHPDYPGVYVFHTFCSAEIQSVLDLEELVAA
GRVPSAGYENIVTKIGHSTEAKSHLGDALRAARHLFSQLPSEVKHRRIFLFTNDVNPH
RGDEELLQKCAIQIESLSSGGVGLVCYDMSPTALPSPAASTQSGEAAAMSERWATQFG
GVDEFWSALTGAVPKATNSSRVGTIDIVHLNSDDSVMLGALSTAVRWRTHPRGASQTT
TLTIGVGAKGSALPRLTVGMYFPMTNAQRRLSKWLDGRAGEMVMLRQRAVGVGAPILL
PCHAGKSGAPPAHISTGQLGKSRPSTGPGTCPNVPLSCKKHIAEVVGAGLTLGFSIIC
FKNADDVLHPQYVLGKSCVLHPDPQDGSDGSLRLFIRLARALKEQRKVAMAQHITRYA
TPPRLVALVPPGLNGDHVDTASFPVMHGLGLYVVPLSYADDVRTTPRSPLFGDATKPM
ERDIALAQRLLATLPSKYNVNVVPNPALELRYKAIESIVQQTQQQSGLSKEIPGVAPF
SQLSKAVDRTWGDRGAMAKYGSVFDEFKAVLLPSYNRDETCGPKTKAARNVKARAELA
DEDIKGVERIISAVESAFRFQSMGMLTEPQLKEYLRVMEGGAFGVRRNPDIIQAVISQ
LQGDI

KU70 DNA sequence (accession number in NCBI nucleotide database = AJ307890)

aagccgctacgcaatccaacagcaaatatgagcgcggaggagtacgagcggtgggtcagt
gagcgcgacgcactgtttgcttcagctgatggtcgccaataaacgtgcggggagaggggc
ttcgagaacggggaagtgtacggctaaataaatatgtatcatcggttggtttttcccttt
tctccttttgttcccgtgagtttgtttttatttagttttatgttagaagggttcacgat
cagtggggtttccgtgtttcatttcttcttttctccgcgaagagttagattagaattatc
cctttatttttttttcgaccgtccccttttttcttgttcctctgctacgtgcttttctat
ttcattgtaggaacgacgtctggcggtaagccgagtatgtatatgtttctgttggggagg
cggggttcaacatgaagcaaagagacccgcacttcgatgtactgccaataa**ccgcaacag**
cgcacacaaccaaagcacgtattcactgccctatttattgactaatgtttgctgacttcc
ttccttctattgttcgcggcatgccttcgtttatcattttgaacaatcgaaaagttattt
ttaccgtttttcttcccttcacttttttgttcgttcattcatccaatcatttgtattact
tactaagcttgtgaaggaggggtgcgccgggttggaccc***atg***gcgcactatgatgagtgg
ctcatgtcagttgacgggctcgatttccacgacgaggatgacagctattttaatgaaact
tatgcaacggaggatgagccttctctccggacgggtagtgtctcttctctagaccagttt
gatgtggtgctgtgcctcgtcgactttcaacaacgcatgttccaaggtgttgttgcgcat
gattgcaagaacccatttgccggtggagcagaacaggtatacgacgatatcaccgaggtg
aaagtgaaagaaacccccacagtgttcgaaaag**gcaatctgctgcgttcagcaa**ctgtac
aaggacaaggggatatctgacagcaacgatatggttgcgttggtgctgtacaatacaagg
gaatgcacacatcctgattaccctggtgtatacgtctttcatacttttgctcagcggaa

atacaatctgtactggatttggaggaacttgtagctgcggggagggtaccgtcagcggga
tatgaaaatattgtgacgaagattggccattcaactgaagccaaatctcatctcggggat
gcgctgcgggcagcccggcatttgttttcccagctgccttccgaggtcaaacatcgtcga
atttttctcttcacaaacgacgtgaacccacatcgcggtgacgaggagcttttgcagaaa
tgtgctatccaaatcgagagcttgtcctctgggggagtgggcttggtgtgctatgatatg
agccccactgcccttccgtc**acctgcggcttcaacgcaat**cgggagaagcggctgcaatg
agtgagcgttgggcaacacagttcggcggtgtggacgagttttggagtgctcttacgggt
gcggtaccgaaggcaacaaacagttcacgggttggcactattgatattgttcacctgaat
agcgacgatagcgttatgctgggagcgctctcaacggcagtccgctggcgcacacaccca
cggggagcctctcaaacgacaacactaactattggtgtaggcgcaaagggctcggcacta
ccaagactgacagtgggtatgtatttcccgatgacaaatgcgcagcgccgattatcgaag
tggttggatggacgcgcaggtgaaatggtaatgcttcggcagcgtgcggtgggagtgggg
gcccctattttgctgccttgccacgcggggaagagtggtgc**gccacctgcacatatcagt**
acggggcagttgggcaagtcacgtccaagcacgggaccggggacgtgcccaaacgttcct
ctaagttgtaaaaagcatatcgccgaggtcgttggggcgggattaactctcggcttttcc
atcatatgcttcaaaaacgctgatgacgttctgcaccctcagtacgtcctggggaagtcg
tgcgtcctccacccagatccgcaagacggtagcgatggctcgcttcgacttttcattcgg
ttggctcgagcgctgaaggagcagcggaaggttgcgatggcgcagcacattacgcgttac
gccactcctcctcgtcttgtcgcacttgtgccgcccgggctcaacggtgatcatgtggac
acggctagtttccctgtgatgcacggtcttggtttgtacgttgtgcctctctcgta**tgct**
gacgacgttcgtaccaccccctcgatccccctctcttcggcgatgcaacgaaacctatggaa
cgtgatattgctcttgcacagcgactgctggccacgctgccgtctaagtacaatgtgaac
gtagttcccaaccctgctttggagctacgctataaagctattgagagtattgtgcaacag
acacagcagcagagtgggttgtcaaaggaaattccaggggtagctccgttctcgcaactg
tcaaaggcagttgatcggacgtggggtgatcgtggggctatggcaaaatacggctctgtg
tttgatgagtttaaggcggtattgctcccttcgtacaaccgagacgaaacgtgtggtcct
aagacgaaagcggcgcggaatgtgaaggcgcgagcagagttggctgacgaggacatcaaa
ggtgttgaaagaataa**tatctgccgtggagtcggcgt**ttcggttccagtctatggggatg
ctgactgaaccgcagttgaaggagtatttgcgcgtgatggagggcggtgcctttggagtc
cggagaaatccggatatcattcaagctgtgatatctcaattgcaaggcgacatt*tag*gtg
tggtggggcttcacctttttttctcgtttgcctatattttcgctttggtttcctcctcggt
tgcgctcttccgccgtcgctcggcggggagggggcgtatataaccattgcgagtgcgcg
ttggcaacggggttggaggtctctttatgacactgtgcggggattcttcgctgaggaaga
taccaaagagataagaaggaaaacaaaacaaatacgagttaaacattgtactaacaatgt
gcagttctgatgttgcatgtgtccattccatggactcactaccacatttacccctttcttcc
tttctccctttttaccagaatatttaagggattattattaaggcttgtgctgtttggctc
ggttgagggtgcagtgtcagtgcgacgtttgccgccacctcttgttttatgtgcaggagt
gccagcgaaaggagagtggtagttggtgtcaagtgacggcacaaaggaaataaagtagag
aatcccactgcaaggggttgttgcggatatacataaaaataactaaggtagcatttatt
ttccacggcaaggtttctgttacattaggtaaagcaatgacacaacctgcccaccttcca
ccgcgcccgagtcacacgtttagcatgcgtacagttgagaagttacaccagcgttttctt
gggagaagcggtacaactacaccagcgcaaaatatgtcgaacattgttgaggccatgcgc
caaaaccttttgcgcatcgaaggtgttcttagttatcgcttggatcaccccagcacggca
ctgaccgacacgagcgag

Appendix 2 Amino acid and DNA sequence for Ku80/*KU80*

DNA sequence was obtained following double pass sequencing of ILTat1.2 genomic lambda clones containing the *KU80* open reading frame. Sequencing oligonucleotides are in bold for KU70U1-U4 and the reverse complementary sequences for KU70S1-S5 are in red (see section

2.5.3). T3 and T7 oligonucleotides were also used for sequencing. Start codons are represented by black bold, italicised, underlined text and stop codons by red bold, italicised, underlined text.

Ku80 amino acid sequence

MAFRTSTIFALDVNCSISSLAQAVEFCRLSVLKTMCPSSYDEVA
LVVAGGCRSYSGGTSTSVSTCGSAAQLSAPCLPAPPSVEFITILHRILRDREEQQGVG
SSANFIETLTLCVEVFNLKKRRKQFREVLYLLTDAHAEVVRKSVFRDVLNALRARGVT
LIVVGIDFSQVTGEGSLLPGESAPSLTSNRVKVDNETVLYTLCKALGNDSRVITLRDA
LVSAAQLVCRKVRSLAQKTVFTIGEVRLATNVLGKVRRMHVPANRTPAPTERVPGSRK
AAHSQLFDEETLRLEGGEEDSQKTCLETPVNKPTASSSNRLSEATKIRGSRGIDAVAC
IPQNQVPVHVLVVDTSYMIAPLDDDPVGTRAFRSIVKALAAQDSALVVRYVRSTDGNP
NMFLCVPLTTGDEDVLFMSRLPFMEEIRCLRFPTLQEMEMECDIPEADAQREQELVSS
IVEEMTVGEEVLHPHRVLNPFVQQYYATQRAMVNRWCSRTAENPLSAHEDSDIKLALL
PQLRGISAGFASPGCKVEQLVSRTREKLETCCRVFAYVPRTTSHSPERKVLWERIPAG
QPPVSETVENTEVASPAAPSTLQCHSGSGSSVTCRPVCANPPHDVVSFIQDKRTMFHP
VAHPLGSGSS

KU80 DNA sequence (accession number in NCBI nucleotide database = AJ311845)

ttcatccgtcttcttcctcggtatgcttgcatgtgctcatttattattaccttcacacgg
tgacatctgatatcaatctggaaaggtacacaacagtttacaccacagggaaccgctg**at**
tgctgctcgcgtagggaactgaaagaaagtataataaaacaaccaagaaaaaaagaaagt
gtgtgtgtgtaaggacgaaaagattatttaacggtaaaccttccaat***atg***gctttccgta
cctcaactattttcgccctagacgtaaactgctcgatttcttcacttgcgcaggctgtgg
agttctgtcgcctttccgtactgaagaccatgtgcccttcttcatatgatgaggtcgctt
tggttgtggctggtggttgcagaagctacagcgggggcacatcaacgagtgttagtactt
gtgggagcgcagctcagctttccgccccatgtttgccggctccaccatctgttgagttca
tcacaatactgcaccgaatcctgcgtgaccggg**aggagcaacaaggtgtggga**tccagtg
caaatttcattgaaactcttactctatgtgtagaggtgtttaacttgaagaaacgaagaa
agcaattcagggaggttttgtatctcttaacggatgcgcatgctgaggtggtcaggaaga
gcgtcttccgcgatgtgctgaacgctctgcgagcgcggggtgtgacacttattgttgtgg
ggattgatttctcccaagtgacaggcgaaggctccctgcttcccggtgaatcagccccat
cactcactagtaaccgtgtgaaggttgataatgaaactgtactgtacaccttgtgtaagg
ctttgggtaacgacagtcgagtaataacattacgtgatgccctggtgagcgctgcgcagt
tggtgtgccgaaaagtacgcagccttgcacaaaagacggtattcacaatcggtgaagtgc
gct**tagctacgaacgtgctcggt**aaagtgagacgaatgcatgtaccagcaaatcggacac
cagcaccgactgagcgagtgccagggagtagaaaggcagctcactcgcagctgtttgacg
aagaaactctccgattagaaggcggggaagaggacagtcagaaaacgtgcctggagactc
ccgtaaacaaaccgaccgcctccagcagcaatcgcctctctgaagctacaaagatacgcg
ggtcacggggtattgatgctgttgcttgcattccacaaaaccaggttcctgttcatgtgc
tggtggtcgatacatcgtacatgatagcaccgctagatgatgaccccgttgggacgaggg
cgttccgaagcattgtaaaggcgctggcagctcaagattcagccttggtggtgcgttatg
tccgatcgactgatggaaaccccaacatgttcctctgcgtgccgtt**aacgacaggtgacg**
aagatgtacttttcatgtccagattacctttcatggaagagatccgctgtctgcgctttc
ccactctacaagagatggaaatggaatgcgacattccggaggccgatgcacagagagaac
aggagcttgtttcatcaattgtagaggaaatgacggtaggtgaagaggtgcttcatccac
atagagtgcttaatcccttcgtgcagcaatactacgccacgcagcgcgctatggtgaacc
gctggtgttcccgaacggcggaaaaccctctcagcgcccatgaggattccgatataaaac
tcgcacttctcccacaactgaggggaatttccgcagggttcgcctcgccaggatgcaaag

tcgagcaacttgtttctcgtacccgtgagaagcttgaaacgtgttgtcgcgttttcgcgt
atgtaccgaggacgaccagtcattctccagagcgcaaagttctatgggagcgcattccag
ctgggcaaccgccggtatctgaaacagttgaaaatacggaagtggcatcacctgcggctc
cttctactctgcagtgtcactcgggaagcggatcctccgttacatgccggcctgtttgtg
ctaatccaccacatgatgtggttagtttcatacaagacaagagaactatgtttcatcctg
tcgcacaccctctggggagcggctcgtcgtaatgcgtggacagtgagtttaacagcagtg
ctctgtagatattttcctaatgggaataaataaataaataaagcaaagataaaacaagac
aagacaagagaacgagagaacagcactctcttgagaaacaaagtcacatataaatatatc
gtactattcttgtatcatgataaaaaagtgggtttggtgctcattcataaattttctttt
ttgcttcatttccccaccctcctcttacatgcgacatttaaaatgcacgtgacgtgccga
gttatttccaccaaagtcaagagtgatggatcttgagagcacgaatgaaaaaggaaaaaa
aataataataataaaagtgaagctgcgcgaaaacacagacgaagcgaaagggcacgtgcg
agtgagaaagatg

REFERENCES

Alarcon, C.M., Pedram, M. and Donelson, J.E. (1999) Leaky transcription of variant surface glycoprotein gene expression sites in bloodstream African trypanosomes. *J.Biol.Chem.* **274**, 16884-16893.

Alexandre, S., Paindavoine, P., Hanocq-Quertier, J., Paturiaux-Hanocq, F., Tebabi, P. and Pays, E. (1996) Families of adenylate cyclase genes in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **77**, 173-182.

Allen, G. and Gurnett, L.P. (1983) Locations of the 6 disulfide bonds in a variant surface glycoprotein (VSG 117) from *Trypanosoma brucei*. *Biochem.J.* **209**, 481-487.

Alsford, S., Wickstead, B., Ersfeld, K. and Gull, K. (2001) Diversity and dynamics of the minichromosomal karyotype in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **113**, 79-88.

Andrulis, A.D., Meiman, A.M., Zapulla, D.C. and Sternglanz, R. (1998) Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* **394**, 592-595.

Aravind, L. and Koonin, E.V. (2000) SAP - a putative DNA-binding motif involved in chromosomal organization. *Trends Biochem.Sci.* **25**, 112-114.

Barber, A.F. and Kamper, S.M. (1993) The importance of mosaic genes to trypanosome survival. *Parasitol.Today* **9**, 63-66.

Barbet, A.F. (1985) Conservation of structure detected in 2 trypanosome surface glycoproteins by amino-acid sequence alignment. *Mol.Biochem.Parasitol.* **14**, 175-185.

Barry, J.D. (1997) The relative significance of mechanisms of antigenic variation in African trypanosomes. *Parasitol.Today* **13**, 212-218.

Barry, J.D., Graham, S.V., Fotheringham, M., Graham, V.S., Kobryn, K. and Wymer, B. (1998) VSG gene control and infectivity strategy of metatcyclic stage *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **91**, 93-105.

Barry, J.D. and McCulloch, R. (2001) Antigenic variation in trypanosomes: Enhanced phenotypic variation in a eukaryotic parasite. *Advances in Parasitology.* **49**, 1-70.

Baumann,P. and Cech,T.R. (2000) Protection of telomeres by the Ku protein in fission yeast. *Mol.Biol.Cell* **11**, 3265-3275.

Baur, J.A., Zou, Y., Shay, J.W., and Wright, W.E. (2001) Telomere position effect in human cells. *Science* **292**, 2075-2077.

Bennet, R.A.O., Gu, X. and Povirk, L.F. (1996) Construction of a vector containing a site-specific DNA double-strand break with 3'-phosphoglycolate termini and analysis of the products of end-joining in CV-1 cells. *Int.J.Radiat.Biol.* **70**, 623-636.

Benson, F.E., Baumann, P. and West, S.C. (1998) Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. *Nature* **391**, 401-404.

Blasco, M.A., Lee, H.W., Hande, M.P., Samper, E., Lansdorp, P.M., DePinho, R.A., Greider, C.W. (1997) Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. *Cell* **91**, 25-34.

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

Borst, P. (1991) Transferrin receptor, antigenic variation and the prospect of a trypanosome vaccine. *Trends Genet.* **7**, 307-309.

Borst, P., Bitter, W., Blundell, P.A., Chaves, I., Cross, M., Gerrits, H., van Leeuwen, F., McCulloch, R., Taylor, M. and Rudenko, G. (1998) Control of VSG gene expression sites in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **91**, 67-76.

Borst, P. and Fairlamb, A.H. (1998) Surface receptors and transporters of *Trypanosoma brucei*. *Annu.Rev.Microbiol.* **52**, 745-778.

Boulton, S.J. and Jackson, S.P. (1996) Saccharomyces cerevisiae Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error- prone DNA repair pathways. *EMBO J.* **15**, 5093-5103.

Boulton, S.J. and Jackson, S.P. (1998) Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing. *EMBO J.* **17**, 1819-1828.

Broccoli, D., Smogorzewska, A., Chong, L. and de Lange, T. (1997) Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat.Genet.* **17**, 231-235.

Brun, R. and Schonenberger, M. (1979) Cultivation and *in vitro* cloning of procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. *Acta Trop.* **36**, 289-292.

Campbell, D.A., Vanbree, M.P. and Boothroyd, J.C. (1984) The 5'-limit of transposition and upstream barren region of a trypanosome VSG gene – tandem 76 base-pair repeats flanking $(TAA)_{90}$. *Nucleic Acids Res.* **12**, 2759-2774.

Carrington, M. and Boothroyd, J. (1996) Implications of conserved structural motifs in disparate trypanosome surface proteins. *Mol.Biochem.Parasitol.* **81**, 119-126.

Carrington, M., Miller, N., Blum, M., Roditi, I., Wiley, D. and Turner, M. (1991) Variant specific glycoprotein of *Trypanosoma brucei* consists of 2 domains each having an independently conserved pattern of cysteine residues. *J.Mol.Biol.* **221**, 823-835.

Cary, R.B., Peterson, S.R., Wang, J.T., Bear, D.G., Bradbury, E.M. and Chen, D.J. (1997) DNA looping by Ku and the DNA-dependent protein kinase. *Proc.Natl.Acad.Sci.USA* **1997**, 4267-4272.

Chen, C. and Kolodner, R.D. (1999) Gross chromosomal rearrangements in *Saccharomyces cerevisiae* replication and recombination deficient mutants. *Nat.Genet.* **23**, 81-85.

Chung, H.M., Lee, M.G.S. and Van der Ploeg, L.H.T. (1992) RNA polymerase I-mediated protein-coding gene-expression in *Trypanosoma brucei. Parasitol.Today* **8**, 414-418.

Chung, H.M.M., Shea, C., Fields, S., Taub, R.N., Van der Ploeg, L.H.T. (1990) Architectural organization in the interphase nucleus of the protozoan *Trypanosoma brucei*: location of telomeres and mini-chromosomes. *EMBO J.* **9**, 2611-2619.

Cockell, M. and Gasser, S.M. (1999) Nuclear compartments and gene regulation. *Curr.Opin.Genet.Dev.* **9**, 199-205.

Collins, A.R. (1993) Mutant rodent cell lines sensitive to ultraviolet light, ionizing radiation and cross-linking agents: a comprehensive survey of genetic and biochemical characteristics. *Mut.Res.* **293**, 99-118.

Coppens, I., Baudhuin, P., Opperdoes, F.R. and Courtoy, P.J. (1998) Receptors for the host low density lipoproteins on the hemoflagellate *Trypanosoma brucei* – purification and involvement in the growth of the parasite. *Proc.Natl.Acad.Sci.USA* **85**, 6753-6757.

Critchlow, S.E., and Jackson, S.P. (1998) DNA end-joining: from yeast to man. *TIBS reviews.* **23**, 394-398.

Cross, G.A.M. (1975) Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei. Parasitology* **71**, 393-417.

Cross, G.A.M. (1990) Glycolipid anchoring of plasma membrane proteins. *Annu.Rev.Cell.Biol.* **6**, 1-39.

Cully, D.F., Ip, H.S. and Cross, G.A.M. (1985) Coordinate transcription of variant surface glycoprotein genes and an expression site associated gene family in *Trypanosoma brucei. Cell* **42**, 173-182.

Delauw, M.F., Laurent, M., Paindavoine, P., Aerts, D., Pays, E., Le Ray, D. and Steinert, M. (1987) Characterisation of genes coding for 2 major metacyclic surface antigens in *Trypanosoma brucei. Mol.Biochem.Parasitol.* **23**, 9-17.

de Bruin, D., Kantrow, S.M., Liberatore, R.A. and Zakian, V.A. (2000) Telomere folding is required for the stable maintenance of telomere position effects in yeast. *Mol.Cell.Biol.* **20**, 7991-8000.

de Bruin, D., Zaman, Z., Liberatore, R.A. and Ptashne, M. (2001) Telomere looping permits gene activation by a downstream UAS in yeast. *Nature* **209**, 109-113.

Deitsch, K.W., Moxon, E.R., and Wellems, T.E. (1997) Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiol.Mol.Biol.Rev.* **61**, 281-294.

DeLange, T., Kooter, J.M., Michels, P.A.M. and Borst, P. (1983) Telomere conversion in trypanosomes. *Nucleic Acids Res.* **11**, 8149-8165.

Diede, S.J. and Gottschling, E. (1999) Telomerase-mediated telomere addition *in vivo* requires DNA primase and DNA polymerases α and δ. *Cell* **99**, 723-733.

Donelson, J.E. (1995) Mechanisms of antigenic variation in *Borrelia hermsii* and African trypanosomes. *J.Biol.Chem.* **270**, 7783-7786.

Donelson, J.E., Hill, K.L., and Elsayed, N.A. (1998) Multiple mechanisms of immune evasion by African trypanosomes. *Mol.Biochem.Parasitol.* **91**, 51-66.

Dynan, W.S. and Yoo, S. (1998) Interaction of Ku proteinand DNA-dependent protein kinase catalytic subunit with nucleic acids. *Nucleic Acids Res.* **26**, 1551-1559.

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

Ersfeld, K. and Gull, K. (1997) Partitioning of large and minichromosomes in Trypanosoma brucci. *Science* **276**, 611-4.

Featherstone, C. and Jackson, S.P. (1999) Ku, a DNA repair protein with multiple cellular functions? *Mutation Research-DNA Repair* **434**, 3-15.

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

Fishman-Lobell, J., Rudin, N. and Haber, J.E. (1992) Two alternative pathways of double-stranded break repair that are kinetically separable and independently modulated. *Mol.Cell.Biol.* **12**, 1292-1303.

Florent, I., Baltz, T., Raibaud, A. and Eisen, H. (1987) On the role of repeated sequences 5' to variant surface glycoprotein genes in African trypanosomes. *Gene* **53**, 55-62.

Frank, K.M., Sekiguchi, J.M., Seidl, K.J., Swat, W., Rathbun, G.A., Cheng, H.L., Davidson, L., Kangaloo, L. and Alt, F.W. (1998) Late embryonic lethality and impaired V(D)J recombination in cells lacking DNA ligase IV. *Nature* **396**, 173-177.

Gao, Y.J., Chaudhuri, J., Zhu, C.M., Davidson, L., Weaver, D.T. and Alt, F.W. A (1998) Targeted DNA PK_{cs} null mutation reveals DNA PK-independent functions for Ku in V(D)J recombination. *Immunity* **9**, 367-376.

Galy, V., Olivo-Marin, J.C., Scherthan, H., Doye, V., Rascalou, N. and Nehrbass, U. (2000) Nuclear pore complexes in the organization of silent telomeric chromatin. *Nature* **403**, 108-112.

Gell, D. and Jackson, S.P. (1999) Mapping of protein-protein interactions within the DNA-dependent protein kinase complex. *Nucleic Acids Res.* **27**, 3494-3502.

Gotta, M., LaRoche, T., Formenton, A., Maillet, L., Scherthan, H. and Gasser, S.M. (1996) The clustering of telomeres and colocalization with Rapl, Sir3, and Sir4 proteins in wild-type *Saccharomyces cerevisiae*. *J.Cell.Biol.* **134**, 1349-1363.

Gottesdiener, K.M. (1994) A new VSG expression site-associated gene (ESAG) in the promoter region of *Trypanosoma brucei* encodes a protein with 10 potential transmembrane domains. *Mol.Biochem.Parasitol.* **63**, 143-151.

Gottschling, D.E., Aparicio, O.M., Billington, B.L. and Zakian, V.A. (1990) Position effect at *Saccharomyces cerevisiae* telomeres – reversible repression of pol II transcription. *Cell* **63**, 751-762.

Grab, D.J., Russo, D., Naessens, J. and Verjee, Y. (1992) Transferrin binding proteins in *Trypanosoma brucei*. *FASEB Journal* **6**, A1900.

Graham, V.S. and Barry, J.D. (1996) Is point mutagenisis a mechanism for antigenic variation in *Trypanosoma brucei? Mol.Biochem.Parasitol.* **79**, 35-45.

Gravel, S., Larrivee, M., Labrecque, P. and Wellinger, R.J. (1998) Yeast Ku as a regulator of chromosomal DNA end structure. *Science* **280**, 741-744.

Grawunder, U., Wilm, M., Wu, X.T., Kulesza, P., Wilson, T.E., Mann, M. and Lieber, M.R. (1997) Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* **338**, 492-495.

Griffith, J.D., Corneau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H. and de Lange, T. (1999) Mammalian telomeres end in a large duplex loop. *Cell* **97**, 503-514.

Grunstein, M. (1997) Molecular model for telomeric heterochromatin in yeast. *Curr.Opin.Cell Biol.* **9**, 383-387.

Haber, J.E. (2000) Partners and pathways repairing a double-strand break. *Trends Genet.* **16**, 259-264.

Haber, J.E. (1998) A locus control region in yeast that regulates recombination. *Trends Genet.* **14**, 317-321.

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

Hoek, M. and Cross, G.A. (2001) Expression-site-associated-gene-8 (ESAG8) is not required for regulation of the VSG expression site in Trypanosoma brucei. *Mol.Biochem.Parasitol.* **117**, 211-215.

Hoek, M., Engstler, M. and Cross, G.A. (2000) Expression-site-associated gene 8 (ESAG8) of Trypanosoma brucei is apparently essential and accumulates in the nucleolus. *J.Cell Sci.* **113**, 3959-3968.

Holder, A.A. and Cross, G.A.M. (1981) Glycopeptides from variant surface glycoproteins of *Trypanosoma brucei* – C-terminal location of antigenically cross-reacting carbohydrate moieties. *Mol.Biochem.Parasitol.* **2**, 135-150.

Hope, M., MacLeod, A., Leech, V., Melville, S., Sasse, J., Tait, A. and Turner, C.M. (1999) Analysis of ploidy (in megabase chromosomes) in Trypanosoma brucei after genetic exchange. *Mol.Biochem.Parasitol.* **104**, 1-9.

Horn, D. and Cross, G.A.M. (1995) A developmentally-regulated position effect at a telomeric locus in *Trypanosoma brucei*. *Cell* **83**, 555-561.

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

Horn,D., Spence,C., and Ingram,A.K. (2000) Telomere maintenance and length regulation in Trypanosoma brucei. *EMBO J.* **19**, 2332-2339.

Johnson, J.G. and Cross, G.A.M. (1979) Selective cleavage of variant surface glycoproteins from *Trypanosoma brucei*. *Biochem.J.* **178**, 689-697.

Kamper, S.M. and Barbet, A.F. (1992) Surface epitope variation via mosaic gene formation is potential key to long-term survival of *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **53**, 22-44.

Kass-Eisler, A. and Greider, C.W. (2000) Recombination in telomere-length maintenance. *Trends Biochem.Sci.* **25**, 200-206.

Klar, A.J., Ivanova, A.V., Dalgaard, J.Z., Bonaduce, M.J. and Grewal, S.I. (1998) Multiple epigenetic events regulate mating-type switching of fission yeast. *Novartis Found Symp.* **214**, 87-99.

Kramer, K.M., Brock, J.A., Bloom, K., Moore, J.K. and Haber, J.E. (1994) Two different types of double-strand breaks in *Saccharomyces cerevisiae* are repaired by similar

RAD52-independent, nonhomologous recombination events. *Mol.Cell.Biol.* **14**, 1293-1301.

LaRoche, T., Martin, S.G., Gotta, M., Gorham, H.C., Pryde, F.V., Louis, E.J. and Gasser, S.M. (1998) Mutation of yeast Ku genes disrupts the subnuclear organization of telomeres. *Curr.Biol.* **8**, 653-656.

Laufer, G., Schaaf, G., Bollgonn, S. and Gunzl, A. (1999) *In vitro* analysis of alpha-amantin-resistant transcription from the rRNA, procyclic acidic repetitive protein, and variant surface glycoprotein gene promoters in *Trypanosoma brucei. Mol.Cell.Biol.* **19**, 5466-5473.

Le, S., Moore, J.K., Haber, J.E. and Greider, C.W. (1999) *RAD50* and *RAD51* define two different pathways that collaborate to maintain telomeres in the absence of telomerase. *Genetics* **152**, 143-152.

Lee, M.G.S. and Van der Ploeg, L.H.T. (1997) Transcription of protein-coding genes in trypanosomes by RNA polymerase I. *Annu.Rev.Microbiol.* **5**, 463-489.

Lee, M.G. and Van der Ploeg, L.H.T. (1987) Frequent independent duplicative transpositions activate a single VSG gene. *Mol.Cell.Biol.* **7**, 357-364.

Lee, S.E., Paques, F., Sylvan, J. and Haber, J.E. (1999) Role of yeast *SIR* genes and mating type in directing DNA double-strand breaks to homologous and non-homologous repair paths. *Curr.Biol.* **9**, 767-770.

Lenardo, M.J., Esser, K.M., Moon, A.M., Van der Ploeg, L.H.T. and Donelson, J.E. (1984) Characterization of the genes specifying two metacyclic variable antigen types in *Trypanosoma brucei rhodesiense. Proc.Natl.Acad.Sci.USA* **81**, 6642-6646.

Lengauer, C., Kinzler, K.W. and Vogelstein, B. (1998) Genetic instabilities in human cancers. *Nature* **396**, 643-649.

Lieber, M.R. (1999) The biochemistry and biological significance of nonhomologous DNA end joining: an essential repair process in multicellular eukaryotes. *Genes to Cells* **4**, 77-85.

Liu, A.Y.C., Van der Ploeg, L.H.T., Rijsewijk, F.A.M. and Borst, P. (1983) The transposition unit of variant surface glycoprotein gene 118 of *Trypanosoma brucei* – presence of repeated elements at its border and absence of promoter-associated sequences. *J.Mol.Biol.* **167**, 57-75.

Longacre, S. and Eisen, H. (1986) Expression of whole and hybrid genes in *Trypanosoma equiperdum* antigenic variation. *EMBO J.* **5**, 1057-1063.

MacLeod, A. (1999) Genetic analysis of *Trypanosoma brucei. Ph.D.thesis, University of Glasgow.*

Mages, G.J., Feldmann, H.M. and Winnacker, E.L. (1996) Involvement of the Saccharomyces cerevisiae HDF1 gene in DNA double-strand break repair and recombination. *J.Biol.Chem.* **271**, 7910-7915.

Majumder, H.K. Boothroyd, J.C. and Weber, H. (1981) Homologous 3'-terminal regions of mRNAs for surface antigens of different antigenic vartiats of *Trypanosoma brucei. Nucleic Acids Res.* **9**, 4745-4753.

Malkova, A., Ivanov, E.L. and Haber, J.E. (1996) Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. *Proc.Natl.Acad.Sci U S A.* **93**, 7131-6.

Manolis, K.G., Nimmo, E.R., Hartsuiker, E., Carr, A.M., Jeggo, P.A., and Allshire, R.C. (2001) Novel functional requirements for non-homologous DNA end joining in Schizosaccharomyces pombe. *EMBO J.* **20**, 210-221.

Martin, S.G., Laroche, T., Suka, N., Grunstein, M. and Gasser, S.M. (1999) Reloclisation of telomeric Ku and Sir proteins in response to DNA double-strand breaks. *Cell* **97**, 609-620.

Matthews, K.R., Shiels, P.G., Graham, S.V., Cowan, C. and Barry, J.D. (1990) Duplicative activation mechanisms of two trypanosome telomeric VSG genes with structurally simple 5' flanks. *Nucleic Acids Res.* **18**, 7219-7227.

McCulloch, R. and Barry, J.D. (1999) A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. *Genes Dev.* **13**, 2875-2888.

McCulloch, R., Rudenko, G. and Borst, P. (1997) Gene conversion mediating antigenic variation in *Trypanosoma brucei* can occur in variant surface glycoprotein expression sites lacking 70 base-pair repeat sequences. *Mol.Cell.Biol.* **17**, 833-843.

Melville, S.E., Leech, V., Gerrard, C.S., Tait, A. and Blackwell, J.M. (1998) The molecular karyotype of the megabase chromosomes of *Trypanosoma brucei* and the assignment of chromosome markers. *Mol.Biochem.Parasitol.* **94**, 155-173.

Miller, E.N., Allan, L.M. and Turner, M.J. (1984) Topological analysis of antigenic determinants on a variant surface glycoprotein of *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **13**, 67-81.

Milne, G.T., Jin, S., Shannon, K.B. and Weaver, D.T. (1996) Mutation in two Ku homologs define a DNA end-joining repair pathway in *Saccharomyces cerevisiae*. *Mol.Cell.Biol.* **16**, 4189-4198.

Morrow, D.M., Connelly, C. and Hieter, P. (1997) "Break copy" duplication: a model for chromosome fragment formation in Saccharomyces cerevisiae. *Genetics* **147**, 371-382.

Moxon, E.R., Rainey, P.B., Nowak, M.A., and Lenski, R.E. (1994) Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr.Biol.* **4**, 24-33.

Munoz-Jordan, J.L., Cross, G.A.M., de Lange, T., and Griffith, J.D. (2001) T-loops at trypanosome telomeres. *EMBO J.* **20**, 579-588.

Murti, K.G. and Prescott, D.M. (1999) Telomeres of polytene chromosomes in a ciliated protozoan terminate in duplex DNA loops. *Proc.Natl.Acad.Sci.U.S.A.* **96**, 14436-14439.

Myung, K., Chen, C. and Kolodner, R.D. (2001) Multiple pathways cooperate in the suppression of genome instability in Saccharomyces cerevisiae. *Nature* **411**, 1073-6.

Namsaraev, E. and Berg, P. (1997) Characterization of strand exchange activity of yeast Rad51 protein. *Mol.Cell.Biol.* **17**, 5359-5368.

Navarro, M. and Cross, G.A.M. (1996) DNA rearrangements associated with multiple consecutive directed antigenic switches in *Trypanosoma brucei*. *Mol.Cell.Biol.* **16**, 3615-3625.

Navarro, M., Cross, G.A.M. and Wirtz, E. (1999) *Trypanosoma brucei* variant surface glycoprotein regulation involves coupled activation/inactivation and chromatin remodeling of expression sites. *EMBO J.* **18**, 2265-2272.

Navarro, M. and Gull, K.A. (2001) A pol I transcriptional body associated with VSG mono-allelic expression in Trypanosoma brucei. *Nature* **414**, 759-763.

Nicolas, A.L., Munz, P.L. and Young, C.S.H. (1995) A modified single-strand annealing model best explains the joining of DNA double-strand breaks in mammalian cells and cell extracts. *Nucleic Acids Res.* **23**, 1036-1043.

Nussenzweig, A., Chen, C., Soares, C.dC, Sanchez, M., Sokol, K., Nussenzweig, M.C. and Li, G.C. (1996) Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature* **382**, 551-555.

Olafson, R.W., Clarke, M.W., Kielland, S.L., Pearson, T.W., Barbet, A.F. and McGuire, T.C. (1984) Amino-terminal sequence homology among variant surface glycoproteins of African trypanosomes. *Mol.Biochem.Paraitol.* **12**, 287-298.

Overath, P., Chaudhri, M., Steverding, D. and Ziegelbauer, K. (1994) Invariant surface proteins in bloodstream forms of *Trypanosoma brucei. Parasitol.Today* **10**, 53-58.

Paindavoine, P., Rolin, S., Van Assel, S., Geuskins, M., Jauniaux, J.C., Dinsart, C., Huet, G. and Pays, E. (1992) A gene from the variant surface glycoprotein expression site encodes one of several transmembrane adenylate cyclases located on the flagellum of *Trypanosoma brucei. Mol.Cell.Biol.* **12**, 1218-1225.

Padilla-Nash, H.M., Heselmeyer-Haddad, K., Wangsa, D., Zhang, H., Ghadimi, B.M., Macville, M., Augustus, M., Schrock, E., Hilgenfeld, E., Ried, T. (2001) Jumping translocations are common in solid tumor cell lines and result in recurrent fusions of whole chromosome arms. *Genes Chromosomes Cancer* **30**, 349-363.

Paques, F. and Haber, J.E. (1999) Multiple pathways of recombination induced by double-strand breaks in *Saccahromyces cerevisiae. Microbiol.Mol.Biol.Reviews* **63**, 349-404.

Pays, E., Delauw, M.F., Laurent, M. and Steinert, M. (1983a) Gene conversion as a mechanism for antigenic variation in trypanosomes. *Cell* **34**, 371-381.

Pays, E., Guyaux, M., Aerts, D., Van Meirvenne, N. and Steinert (1985) Telomeric reciprocal recombination as a possible mechanism for antigenic variation in trypanosomes. *Nature* **316**, 562-564.

Pays, E., Laurent, M., Delinte, K., Van Meirvenne, N., and Steinert, M. (1983) Differential size variations between transcriptionally active and inactive telomeres of *Trypanosoma brucei. Nucleic Acids Res.* **11**, 8137-8147.

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., Van Assel, S., Laurent, M., Dero, B., Michiels, F., Kronenberger, P., Matthyssens, G., Van Meirvenne, N., Le Ray, D. and Steinert, M. (1983) At least two transposed sequences are associated in the expression site of a surface antigen gene in different trypanosome clones. *Cell* **34**, 359-369.

Pedram, M. and Donelson, J.E. (1999) The anatomy and transcription of a monocistronic expression site for a metacyclic variant surface glycoprotein gene in *Trypanosoma brucei*. *J.Biol.Chem.* **274**, 16876-16883.

Ray, A. and Runge, K.W. (1999) The yeast telomere length counting machinery is sensitive to sequences at the telomere-nontelomere junction. *Mol.Cell.Biol.* **19**, 31-45.

Redpath, M.B., Windle, H., Nolan, D., Pays, E., Voorheis, H.P. and Carrington, M. (2000) ESAG11, a new VSG expression site-associated gene from Trypanosoma brucei. *Mol.Biochem.Parasitol.* **111**, 223-228.

Robinson, N.P., Burman, N., Melville, S.E. and Barry, J.D. (1999) Predominance of duplicative VSG gene conversion in antigenic variation in African trypanosomes. *Mol.Cell.Biol.* **19**, 5839-5846.

Roditi, I. and Pearson, T.W. (1990) The procyclin coat of African trypanosomes (or the not-so-naked trypanosome). *Parasitol.Today* **6**, 79-82.

Roth, C., Bringaud, F., Layden, R.E., Baltz, T. and Eisen, H. (1989) Active late-appearing variable surface antigen genes in *Trypanosoma equiperdum* are constructed entirely from pseudogenes. *Proc.Natl.Acad.Sci.USA* **86**, 9375-9379.

Rudenko, G., Chaves, I., Dirksmulder, A. and Borst, P. (1998) Selection for activation of a new variant surface glycoprotein gene expression site in *Trypanosoma brucei* can result in deletion of the old one. *Mol.Biochem.Parasitol.* **95**, 97-109.

Rudenko, G., McCulloch, R., Dirksmulder, A. and Borst, P. (1996) Telomere exchange can be an important mechanism of a variant surface glycoprotein gene switching in *Trypanosoma brucei. Mol.Biochem.Parasitol.* **80**, 65-75.

Rudenko, G., Blundell, P.A., Dirksmulder, A., Kieft, R., and Borst, P. (1995) A ribosomal DNA promoter replacing the promoter of a telomeric VSG gene expression site can be efficiently switched on and off in *Trypanosoma brucei. Cell* **83**, 547-553.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual. 2edn, Cold Spring Harbor Laboratory Press.

Samper, E., Goytisolo, F.A., Slijepcevic, P., van Buul, P.P.W., and Blasco, M.A. (2000) Mammalian Ku86 protein prevents telomeric fusions independently of the length of TTAGGG repeats and the G-strand overhang. *Embo Reports* **1**, 244-252.

Schell, D., Evers, R., Preis, D., Ziegelbauer, K., Kiefer, H., Lottspeich, F., Cornelissen, A.W.C.A. and 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 J.* **10**, 1061-1066.

Scherf, A., Figueiredo, L.M., and Freitas-Junior, L.H. (2001) Plasmodium telomeres: a pathogen's perspective. *Curr.Op.Microbiol.* **4**, 409-414.

Schwartz, J.L. (1989) Monofunctional alkylating agent-induced S phase-dependent DNA damage. *Mutat.Res.* **216**, 111-118.

Shah, J.S., Young, J.R., Kimmel, B.E. Iams, K.P. and Williams, R.O. (1987) The 5'flanking sequencing of a *Trypanosoma brucei* variable surface glycoprotein gene. *Mol.Biochem.Parasitol.* **24**, 163-174.

Sherman, J.M. and Pillus, L. (1997) An uncertain silence. *Trends Genet.* **13**, 308-313.

Shore, D. (1997) Telomere length regulation: getting the measure of chromosome ends. *Biol.Chem.* **378**, 591-597.

Siede, W., Friedl, A.A., Dianova, I., EckardtSchupp, F., and Friedberg, E.C. (1996) The Saccharomyces cerevisiae Ku autoantigen homologue affects radiosensitivity only in the absence of homologous recombination. *Genetics* **142**, 91-102.

Signon, L., Malkova, A., Naylor, M.L., Klein, H. and Haber, J.E. (2001) Genetic requirements for RAD51- and RAD54-independent break-induced replication repair of a chromosomal double-strand break. *Mol.Cell.Biol.* **21**, 2048-2056.

Song, K., Jung, D., Jung, Y., Lee, S.G. and Lee, I. (2000) Interaction of human Ku70 with TRF2. *FEBS Lett.* **481**, 81-85.

Stansel, R.M., de Lange, T. and Griffith, J.D. (2001) T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang. *EMBO J.* **20**, 5532-5540.

Steverding, D., Stierhof, Y.D., Fuchs, H., Tauber, R. and Overath, P. (1995) Transferrin-binding protein complex is the receptor for transferring uptake in *Trypanosoma brucei*. *J.Cell Biol.* **131**, 1173-1182.

Sugawara, N. and Haber, J.E. (1992) Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation. *Mol.Cell.Biol.* **12**, 563-575.

Sung, P. (1994) Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science* **265**, 1241-3.

Sung, P., Robberson, D.L. (1995) DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. *Cell* **82**, 453-461.

Sung, P. and Stratton, S.A. (1996) Yeast Rad51 recombinase mediates polar DNA strand exchange in the absence of ATP hydrolysis. *J.Biol.Chem.* **271**, 27983-27986.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J. and Stahl, F.W. (1983) The double-strand-break repair model for recombination. *Cell* **33**, 25-35.

Tait, A., Turner, C.M., Le Page R.W. and Wells, J.M. (1989) Genetic evidence that metacyclic forms of *Trypanosoma brucei* are diploid. *Mol.Biochem.Parasitol.* **37**, 247-255.

Tait, A., Barry, J.D., Wink, R., Sanderson, A. and Crowe, J.S. (1985) Enzyme variation in *T. brucei* ssp. II. Evidence from *T. b. rhodesiense* being a set of variants of *T. b. brucei*. *Parasitilogy* **90**, 89-100.

Takata, M., Sasaki, M.S., Sonoda, E., Morrison, C., Hashimoto, M., Utsumi, H., Yamaguchi-Iwai, Y., Shinohara, A. and Takeda, S. (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.* **17**, 5497-5508.

Taylor, E.M. and Lehmann, A.R. (1998) Conservation of eukaryotic DNA repair mechanisms. *Int.J.Radiat.Biol.* **74**, 277-286.

Thon, G., Baltz, T., and Eisen, H. (1989) Antigenic diversity by the recombination of pseudogenes. *Genes Dev.* **3**, 1247-1254.

Thon, G., Baltz, T., Giroud, C. and Eisen, H. (1990) Trypanosome variable surface glycoproteins: composite genes and order of expression. *Genes Dev.* **9**, 1374-1383.

Timmers, H.T.M., de Lange, T., Kooter, J.M. and Borst, P. (1987) Coincident multiple activations of the same surface antigen gene in *Trypanosoma brucei*. *J.Mol.Biol.* **194**, 81-90.

Turner, C.M.R., Barry, J.D., Maudlin, I. and Vickerman, K. (1988) An estimate of the size of the metacyclic variable antigen repertoire of *Trypanosoma brucei rhodesiense*. *Parasitology*. **97**, 269-276.

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

van der Burgt, I., Chrzanowska, K.H., Smeets, D. and Weemae, C. (1996) Nijmegen breakage syndrome. *J.Med.Genet.* **33**, 153-156.

Vanhamme, L., Poelvoorder, P., Pays, A., Tebabi, P., Xong, H.V. and Pays, E. (2000) Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. *Mol.Microbiol.* **36**, 328-340.

Vickerman, K. (1985) Developmental cycles and biology of pathogenic trypanosomes. *British Medical Bulletin* **41**, 105-114.

Voelkel-Meiman, K. and Roeder, G.S. (1990) Gene conversion tracts stimulated by HOT1-promoted transcription are long and continuous. *Genetics* **126**, 851-67

Wach, A., Brachat, A., Pohlmann, R. and Philippsen, P. (1994) New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. *Yeast* **10**, 1793-1808.

Walker, J.R., Corpina, R.A., and Goldberg, J. (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412**, 607-614.

Webster, P. and Russel, D.G. (1993) The flagellar pocket of trypanosomatids. *Parasitol.Today* **9**, 201-206.

Weiden, M., Osheim, Y.N., Beyer, A.L. and Van der Ploeg, L.H. (1991) Chromosome structure: DNA nucleotide sequence elements of a subset of the minichromosomes of the protozoan *Trypanosoma brucei*. *Mol.Cell.Biol.* **11**, 3823-3834.

Wu, X., Wu, C. and Haber, J.E. (1997) Rules of donor preference in *Saccharomyces* mating-type gene switching revealed by a competition assay involving two types of recombination. *Genetics* **147**, 399-407.

Zhu, X.D., Kuster, B., Mann, M., Petrini, J.H. and Lange, T. (2000) Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. *Nat.Genet.* **25**, 347-352.

Ziegelbauer, K. and Overath, P. (1990) Surface antigen change during differentiation of *Trypanosoma brucei*. *Biochem.Soc.Trans.* **18**, 731-733.

Zomerdijk, J.C., Kieft, R. and Borst, P. (1992) A ribosomal RNA gene promoter at the telomere of a mini-chromosome in *Trypanosoma brucei*. *Nucleic Acids Res.* **20**, 2725-2734.

Zomerdijk, J.C.B.M., Kieft, R., Shiels, P.G. and Borst, P. (1991) Alpha-amanitin-resistant transcription units in trypanosomes – a comparison of promoter sequences for a VSG gene expression site and for the ribosomal RNA genes. *Nucleic Acids Res.* **19**, 5153-5158.

Zomerdijk, J.C.B.M., Ouellette, M., ten Asbroek, A.L.M.A., Kieft, R., Bommer, A.M.M., Clayton, C.E. and Borst, P. (1990) The promoter for a variant surface glycoprotein gene expression site in *Trypanosoma brucei*. *EMBO J.* **9**, 2791-2801.