RAD51-LIKE GENES IN *TRYPANOSOMA BRUCEI*: A POTENTIAL ROLE IN ANTIGENIC VARIATION?

CHRIS PROUDFOOT
WELLCOME CENTRE FOR MOLECULAR PARASITOLOGY
THE ANDERSON COLLEGE
UNIVERSITY OF GLASGOW

Submitted for the Degree of Doctor of Philosophy
February 2005

ABSTRACT

African trypanosomes, such as *Trypanosoma brucei*, are single celled eukaryotic parasites of mammals and are the causative agents of sleeping sickness (trypanosomiasis) in humans and nagana in cattle. To avoid being eliminated by the host's immune response trypanosomes undergo a process termed antigenic variation. This consists of spontaneous, periodic changes in the Variant Surface Glycoprotein (VSG) species that acts as a protective coat on the surface of the parasite.

Only one gene has been identified thusfar to have a role in VSG switching: *RAD51*. This encodes the eukaryotic homologue of bacterial RecA and archaebacterial RADA, and is central to the catalysis of homologous recombination. *rad51-/- T. brucei* cells show increased DNA damage sensitivity, have an impaired recombination and reduced levels of VSG switching. However, recombination and VSG switching do still occur, meaning that backup pathways must exist for both processes.

The *T. brucei* genomic databases were searched with *T. brucei RAD51*, *Saccharomyces cerevisiae* Rad51 and *Escherichia coli* RecA sequences to define genes with the potential to encode strand exchange proteins. Five further *RAD51*-like genes were found in this way, and three were chosen for further analysis. The first was named *RAD51-3*, which has homology to the *S. cerevisiae* RAD51 co-factor Rad57, as well as *H. sapiens* Rad51C, in BLAST searches. The second, named *DMC1*, encodes a protein that is highly homologous to a meiosis-specific form of RAD51 found in many eukaryotes. Finally, the most distantly related of the *RAD51*-like genes was examined, this was named *RAD51-5* and was identified by searching the *T. brucei* genome with *E. coli* RecA.

Genetic and biochemical analyses of homozygous mutants of the three genes identified no role for *DMC1* in repair, recombination or VSG switching. In contrast, both *RAD51-3* and *RAD51-5* were shown to have roles in repair, recombination and to mediate the re-localisation of RAD51 in following DNA damage. Surprisingly, only *RAD51-3* was shown to have a role in VSG switching.

CONTENTS

List of Figures

LIST OF TABLES

ACKNOWLEDGEMENTS

Firstly, I would like to thank Richard, whose bright idea this was in the first place. The completion of this project owes much to his advice and guidance. I would also like to take the opportunity to thank all the members of the Barry/McCulloch group, past and present, who have provided much help and advice, and were always willing to join me when a visit to the Aragon was required.

Thanks must also go to my family who over the years have provided all to often required financial assistance.

Finally, I would like to thank Becky for her support over the course of this project and for putting up with me throughout.

AUTHOR'S DECLARATION

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

CHRISTOPHER PROUDFOOT

CHAPTER 1

INTRODUCTION

1.1 Introduction

Protozoan parasites cause some of the world's most prevalent and devastating diseases. One such protozoan parasite is *Trypanosoma brucei*, the causative agent of African trypanosomiasis, often called sleeping sickness in humans and nagana in cattle. According to the World Health Organisation it is estimated that over 60 million people are at risk from the disease, with 300,000 to 500,000 cases annually (www.who.int/mediacentre/factsheets/fs259/en/). The numbers of deaths from sleeping sickness is estimated at 50,000 annually, although it is considered that this is an underestimate. *T. brucei* is endemic across sub-Saharan Africa and can cause two clinically distinct forms of the human disease depending on which morphologically indistinguishable sub-species of the parasite is involved. The general symptoms of the early phase of the disease are a general malaise, headaches and bouts of fever (Barrett *et al.*, 2003)(www.who.int/mediacentre/factsheets/fs259/en/). The second phase of the disease occurs after the parasite has crossed the blood-brain barrier invading the CNS. This results in confusion, sensory disturbances, poor coordination and sleep disruption (the symptom which gives the disease its name) (Barrett *et al.*, 2003). The disease is fatal if left untreated and the neurological damage caused after onset of the second phase of the disease cannot be reversed by treatment (www.who.int/mediacentre/ factsheets/fs259/en/). *T. brucei gambiense* is endemic to west and central Africa and causes a chronic infection in which a person can be infected for months or years before symptoms emerge, at which point the disease is at an advanced stage. *T. brucei rhodesiense*, in contrast, is endemic to southern and east Africa and causes a more acute form of the disease whose course from infection to death can be weeks or months. A third sub-species of the parasite, *T. brucei brucei*, which is not human infective, affects wildlife and cattle, which can also act as reservoirs of the *T. b. gambiense* and *T. b. rhodesiense* sub-species (www.who.int/mediacentre/ factsheets/fs259/en/).

African trypanosomes are unicellular, flagellated protozoa of the order Kinetoplastida, named after the highly unusual organisation of DNA (the kinetoplast) found in the single large mitochondrion at the base of the flagellum. A number of further species of *Trypanosoma*, including *T. vivax* and *T. congolense*, also affect wildlife and cattle in Africa. Moreover, some *Trypanosoma* have spread beyond Africa, and can infect organisms other than mammals (Barrett *et al.*, 2003; Stevens and Gibson, 1999).

T. brucei has been shown to be able to undergo genetic exchange, with all evidence suggesting that meiotic division is involved, despite the fact that a haploid stage has not

been defined (Gibson, 2001). The first evidence for genetic exchange in *T. brucei* was the production of hybrid trypanosome strains by the co-transmission of two parental strains through tsetse flies (Jenni *et al.*, 1986). Whether or not this genetic exchange occurs at a significant frequency in natural populations is a matter for debate. However, minisatellite analysis of trypanosome isolates from a number of African locations suggests that frequent genetic exchange may occur in the field (MacLeod *et al.*, 1999).

1.1.1 The life cycle of *T. brucei*

T. brucei proliferates in the bloodstream, capillaries and tissue spaces of its mammalian host and is transmitted between hosts by the tsetse fly (*Glossina* spp.). As a result, the parasites have an intricate life cycle (reviewed in Barry and McCulloch, 2001; Matthews *et al.*, 2004; Matthews, 2005) with several morphologically distinct life cycle stages, each with specific adaptations for growth (Fig 1.1). A major division in the life cycle is clear between the mammalian host and tsetse fly vector stages but life cycle stages are also distinguishable between growth in the mid-gut and salivary gland of the fly. In each life cycle stage replicative forms of the parasite (bloodstream long slender, procyclic and epimastigote) can be detected and appear responsible for the establishment of the parasite infection. Equally, at each point of transmission between stages non-replicative forms (bloodstream short stumpy, mesocyclic and metacyclic) can be found (Barry and McCulloch, 2001; Matthews, 2005).

Population growth in the mammalian host is carried out by the long slender form which rapidly, mitotically divides. The trypanosome evades the host immune system by spontaneous, periodic changes of the variant surface glycoprotein (VSG) present on the external surface of the parasite cell membrane; this is a process termed antigenic variation and is discussed in section 1.3 (reviewed in Barry and McCulloch, 2001). The VSG molecules form a densely packed coat that protects against the host alternative complement system and prevents access of the host acquired immune system to invariant surface molecules. Immune reactions generated against the VSG, which is highly immunogenic, result in destruction of the parasite. However, parasites that have switched to expressing an antigenically distinct VSG species survive and proliferate. This contributes to the fluctuating parasitemia and relapse peaks that are characteristic of trypanosome infections (Barry and McCulloch, 2001; Matthews *et al.*, 2004). Trypanosome populations at parasitemic peaks differentiate into the non-dividing short

stumpy form, in a process thought to be density-dependent (Tyler *et al.*, 2001). The short-stumpy form has a finite life, being destroyed by the host immune system or degenerating over the course of a few days if not transmitted to the tsetse fly (Turner *et al.*, 1995). Differentiation is thought to result from the release of a pheromone-like factor, or catabolite, termed stumpy induction factor (SIF) which, although it remains to be identified, appears to act through a cAMP signalling pathway to induce growth arrest (Vassella *et al.*, 1997). Short stumpy forms have several pre-adaptations for growth in the tsetse fly, including metabolic changes to aid movement from the host bloodstream, where glucose is the main energy source, to the tsetse mid-gut, where proline is the main energy source (Hendriks *et al.*, 2000).

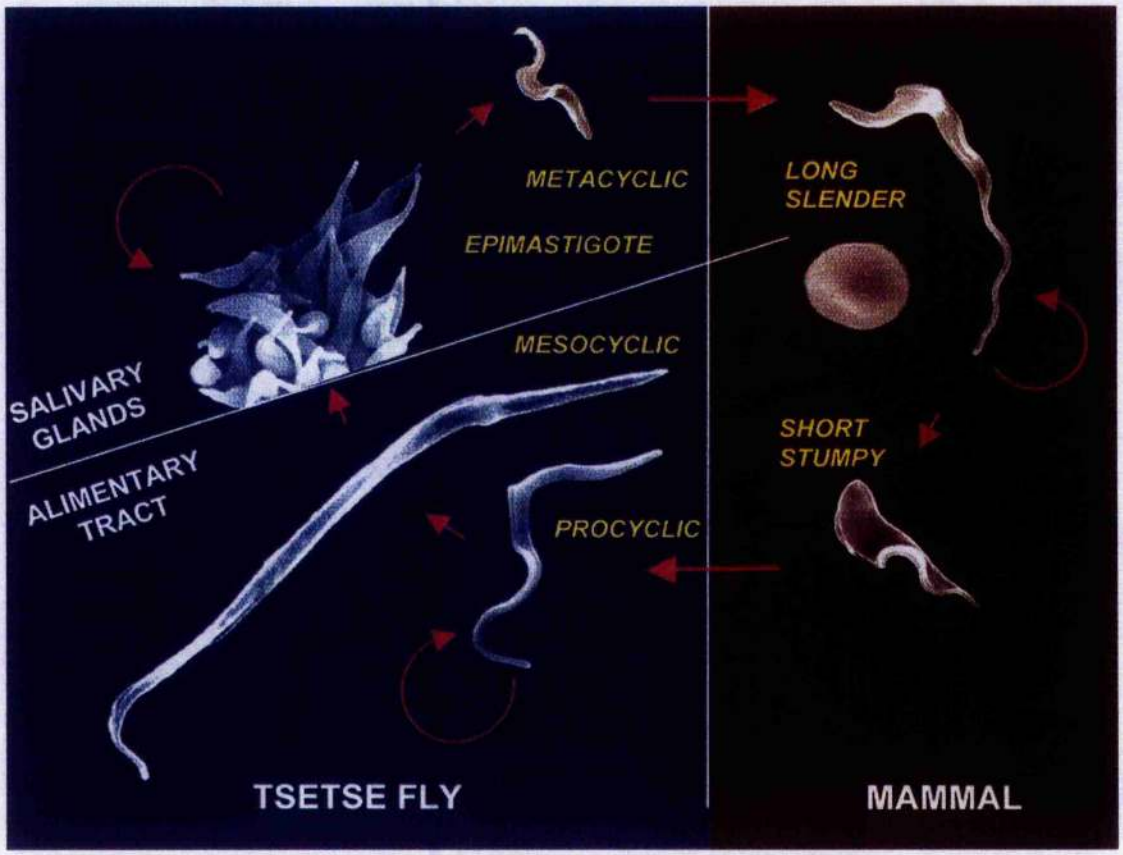


Figure 1.1: The Life-cycle of *T. brucei*. Life-cycle stages are represented as scanning electron microscope images, reproduced to scale. The mammalian stages are shaded red and the tsetse stages shaded blue, circular arrows represent replicative stages. This diagram was taken from Barry & McCulloch (2001).

Upon uptake in the blood meal of the tsetse fly, short stumpy forms differentiate to procyclic forms. Differentiation involves cell lengthening, repositioning of the

kinetoplast and the expression of procyclins (reviewed in Roditi *et al.*, 1998), the major surface molecules in the fly, which replace the VSG coat. Procyclins are thought to protect against digestion by hydrolases in the tsetse midgut and function in tropism (Roditi and Liniger, 2002). Procyclic cells are also released from the cell cycle arrest imposed on the short stumpy form and are able to proliferate.

Despite pre-adaptation to life in the tsetse fly, the majority of ingested trypanosomes do not survive to differentiate (Van den Abbeele *et al.*, 1999). Growth of the surviving population within the tsetse midgut is rapid and reaches a maximum density of 5×10^5 parasites per midgut (Van den Abbeele *et al.*, 1999). During growth of the trypanosome population in the fly there is a migration from the posterior towards the anterior midgut, a migration that coincides with the differentiation to the mesocyclic form. In addition, there are changes in the form of procyclin expressed on the cells surface (Vassella *et al.*, 2001). Mesocyclic cells subsequently enter the foregut and the probiscus. During this movement the cells elongate, replicate their DNA to become 4N and differentiate to become epimastigote cells, the life cycle form that is thought to migrate to the salivary glands.

The division of long mesocyclic cells results in the production of a long and a short epimastigote cell, both of which are 2N (Van den Abbeele *et al.*, 1999). The fate of the long epimastigote cells is unknown, but presumably they fail to differentiate or survive. The short epimastigote cells, however, continue the life cycle by dividing and becoming attached to the epithelial membrane of the salivary glands; it is at this stage that meiotic exchange is most likely to take place (Gibson and Whittington, 1993; Tait *et al.*, 1989). Differentiation of epimastigote cells then occurs to form mammalian-infective metacyclic cells, which are characterised by a cessation of cell division, the gaining of a metacyclic VSG (MVSG) coat and repression of the mitochondrion (Tetley and Vickerman, 1985; Tyler and Engman, 2000).

On entry into the mammalian host via a fly bite trypanosome cells proliferate and disseminate from their initial subcutaneous location (chancre) to the vascular system (Barry and McCulloch, 2001). The MVSG present on the surface of the parasite remains there for the first few days following entry into the host before being replaced with a bloodstream form VSG, and the initiation of antigenic variation.

1.1.2 The genome of *T. brucei*

T. brucei has a haploid nuclear genome size of approximately 35 Mb, varying between isolates by up to 25% (El Sayed *et al.*, 2000). The chromosomes that make up the *T. brucei* genome are categorised into three classes according to size: the megabase chromosomes, the intermediate chromosomes and the mini-chromosomes. The megabase chromosomes are the largest and vary in size from 1 Mb to over 6 Mb. There are 11 of these chromosomes, named I to XI in ascending order of size in the TREU 927/4 strain (Melville *et al.*, 1998). The Megabase chromosomes are diploid, showing mendelian inheritance and constitute approximately 53.4 Mb of DNA (Melville *et al.*, 1998). Only the megabase chromosomes have been shown to be diploid, with the ploidy of the intermediate and mini-chromosomes unknown at present.

In the *T. brucei* genome there are generally between 1 and 5 intermediate chromosomes that range in size from 200 to 900 kb (Ersfeld *et al.*, 1999). Interestingly, none of 401 cDNA and gene clones tested was found to hybridise to the intermediate chromosomes, indicating that they may not contain house keeping genes (Melville *et al.*, 1998).

The mini-chromosomes, present in approximately 100 copies per genome, are linear molecules of 50 to 100 kb in length (Ersfeld *et al.*, 1999). The mini-chromosomes posses the same telomeric repeats as the other larger chromosomes and are composed almost entirely (>90%) of 177 bp repeats (El Sayed *et al.*, 2000). Despite the presence of silent *VSGs* on the telomeres of mini-chromosomes, none have been shown to posses an active *VSG* expression site (discussed below), suggesting that to be activated a mini-chromosomal *VSG* must be duplicated into an active expression site or be involved in a telomere exchange with one (El Sayed *et al.*, 2000; See below).

1.2 Phase and Antigenic variation

Many pathogenic organisms have adopted mechanisms to evade destruction by host defence systems. Phase and Antigenic variation are two of these mechanisms. Phase variation refers to a reversible switch between two distinct states, resulting in many altered phenotypes and behaviours, including evasion of host defences. Antigenic variation, in contrast, refers to the expression of antigenically distinct versions of a functionally conserved moiety, with the specific function of evading acquired immunity. A brief summary of examples of phase and antigenic variation are discussed

below; however, further information can be obtained from the book 'Antigenic Variation' (Craig and Scherf, 2003), and several recent reviews are also available (Allred and Al Khedery, 2004; Borst, 2002; Galinski and Corredor, 2004; Nash, 2002; van der Woude and Baumler, 2004).

Phase variation in bacteria can influence either surface proteins or carbohydrates, and can allow switches between distinct forms, or between on and off expression. A number of mechanisms can account for this process, including gene transcription, promoter inversion, and transcriptional or translational strand slippage. An example of the later is provided by *Helicobacter pylori*, a Gram-negative gastric bacterium present within the stomachs of as much as half of the world population. Phase variation in *H. pylori* is generated during the replication of genes containing homopolymeric repeats, one particular example being those whose products are involved in lipopolysaccharide (LPS) synthesis. DNA polymerase slippage on these C-tracts yields daughter stands with either increased or reduced numbers of C residues, resulting in a reversible frameshift (Appelmelk and Vandenbrouck-Grauls, 2003). This frameshifting results in the fluctuation between functional and non-functional LPS products, causing the generation of a heterogeneous population in terms of LPS expression. Changes that are advantageous to an individual, such as serum resistance, aid survival and proliferation.

Phase variation appears to be confined to bacteria, whereas antigenic variation is found in diverse pathogens, including viruses, bacteria, fungi and protozoan. During host infection of mammals, *Plasmodium falciparum* colonises erythrocytes, causing alterations in their surface morphology which target them for destruction in the spleen. To counter this, the parasites express proteins on the surface of the erythrocyte which cause adherence to the endothelium of the host blood vessels preventing their movement to the spleen. *P. falciparum* achieves this by expression of the *var* genes, which encode the erythrocyte membrane protein 1 (PfEMP1). This also results in the generation of an immune target for the host, so to combat this *P. falciparum* contains approximately 50 *var* genes, of which one is expressed at one time, resulting in the ability to produce PfEMP1 molecules that are antigenically distinct and adhere to different endothelial surface molecules (reviewed by Borst, 2002; Deitsch and Hviid, 2004). Approximately 18% of the parasite population per generation switch *var* gene expression, much faster than previously thought (Gatton *et al.*, 2003). However, recent evidence suggests that variable *var* transition rates also play a part in *P. falciparum* antigenic variation (Horrocks *et al.*, 2004).

Giardia lamblia is a parasite of the gut and evades the host immune system by expression of variant surface protein (VSP) present on the surface of the parasite (reviewed by Nash, 2002). There are approximately 150 VSP species encoded by the *G. lamblia* genome and the switching rate is isolate- and VSP-dependent. The exact mechanism of the regulation of VSP switching is not known. However, it is probably regulated at the transcriptional level (as is var gene switching in *P. falciparum*), either only one VSP is expressed, or all but one repressed, at one time in one cell.

1.3 Antigenic variation in *T. brucei*

Antigenic variation in *T. brucei* is probably the best characterised mechanism for evasion of the host immune system in protozoans and as a result several reviews are available (Barry and McCulloch, 2001; Barry and McCulloch, 2003; Donelson, 2003; Pays *et al.*, 2004; Vanhamme *et al.*, 2001b). Antigenic variation in *T. brucei* involves the spontaneous, periodic switching of the VSG species present on the surface of the parasite. To be expressed a *VSG* gene must occupy a specialised, telomeric transcription unit known as an expression site (ES; see section 1.3.2). There are two types of ES, the bloodstream form ES (BES) for expression of VSG during growth within the bloodstream of the mammalian host, and the metacyclic ES (MES) that allows expression of VSG in the metacyclic cell as a preadaptation for entry in to the mammalian host. Only one ES is active in a single cell at any one time, resulting in a single *VSG* gene being expressed at any one time. Trypanosomes can switch *VSG* expression by one of two apparently distinct mechanisms. They either inactivate transcription from the active ES and activate transcription from a previously inactive ES (*in situ* switch), or move a silent *VSG* into the active ES by DNA recombination. The process of VSG switching is spontaneous and is not dictated by the host immune response. In addition, the switching rates of different trypanosome isolates varies greatly. So-called 'pleomorphic' *T. brucei* lines switch at a rapid rate, around 1×10^{-2} switches/cell/generation (Turner and Barry, 1989), much greater than the background rates of mutation. In contrast 'monomorphic' lines, generated through serial syringe passaging in mice, switch at a much lower rate of around 1×10^{-6} to 1×10^{-7} switches/cell/generation (Lamont *et al.*, 1986). Pleomorphic lines are distinguished from monomorphic lines in that they retain the capacity to differentiate from long-slender to short stumpy bloodstream forms, a process lost in monomorphic cells.

Whether or not there is a connection between this distinction in differentiation capacity and in VSG switch rate has not been experimentally examined.

1.3.1 The VSG species of *T. brucei*

Central to antigenic variation in *T. brucei* is the VSG protein, which forms a densely packed layer on the surface of the parasite comprising approximately 10 million VSG dimers (Blum *et al.*, 1993; Carrington *et al.*, 1991)(Fig. 1.2). The tightly packed VSG species prevent access of the host immune system to the invariant surface molecules, such as receptors for transferrin or lipoproteins and glucose transporters (Borst and Fairlamb, 1998; Pays and Nolan, 1998), allowing the parasite to survive in the bloodstream. To facilitate the process of antigenic variation, trypanosomes contain a large repertoire of distinct VSG species present as predominantly silent genes. Original estimates suggested that *T. brucei* contained around 1000 *VSGs* (Van der Ploeg *et al.*, 1982), and subsequent sequence analysis of the genome of *T. brucei* strain 927 has shown that this to be a reasonable estimate, with approximately 900 *VSGs* currently annotated (L. Marcello and J.D. Barry, pers. comm.).

To date, 877 *VSGs* have been annotated, and all but one are located in sub-telomeric domains. A minority are present in the ESs, and the majority are in silent arrays. Surprisingly, only 35 (4%) of these *VSGs* were found to be functional. Of the remainder, 576 (66%) were found to be full length pseudogenes, 73 (8%) were Atypical and 184 (22%) were shown to be fragments (L. Marcello and J.D. Barry, pers. comm.). The *T. brucei* genome sequence does not include the 100 or so mini-chromosomes, increasing the size of the repertoire, but all the *VSGs* characterised to date on this class of chromosome appear to be functional (Alsford *et al.*, 2001).

The three dimensional structure of an expressed VSG consists of a 350-400 amino acid residue N-terminal region and a 50-100 residue C-terminal region which are separated by a hinge (Carrington *et al.*, 1991). The sequences of different VSG proteins show high levels of divergence, with the levels of identity between two N-terminal domains as low as 20% (Blum *et al.*, 1993). The N-terminal region encodes the antigenic epitopes present on the surface of the parasite, and therefore this level of sequence divergence results in the generation of a large number of immunogenically distinct epitopes. The C-terminal region is attached to the plasma membrane of the trypanosome via a glycophosphatidylinositol (GPI) anchor (Boothroyd *et al.*, 1980). The function of the C-

terminal region remains obscure, however, it has been suggested that it acts to lengthen the VSG, resulting in a thicker coat (Chattopadhyay *et al.*, 2005).

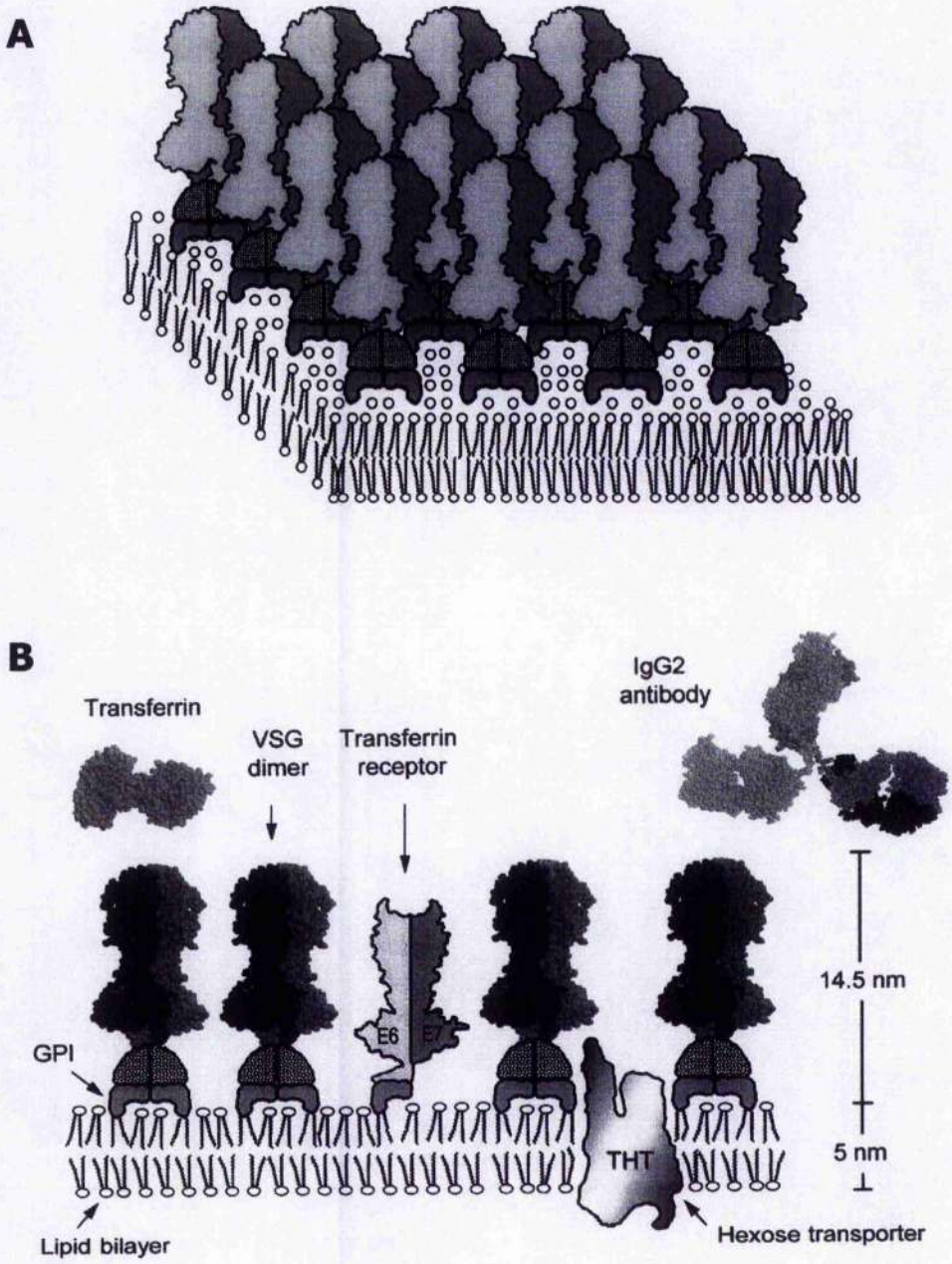


Figure 1.2: The surface of *T. brucei*. (A) A three-dimensional depiction of the tightly packed VSG dimers present on the cell surface of bloodstream form *T. brucei*. (B) A schematic representation of the VSGs and the invariant molecules present on the surface of *T. brucei*. The invariant surface molecules present are the Transferrin receptor, which is encoded by ESAGs 6 and 7, and a Hexose transporter. An IgG2 antibody is shown for size comparison. Figure taken from Borst and Fairlamb (1998).

1.3.2 The expression sites of *T. brucei*

As stated previously (section 1.3), to be expressed a *VSG* gene must occupy either a BES or MES (Fig. 3). The estimated numbers of BES found in the *T. brucei* genome vary from approximately 30 (Navarro and Cross, 1996) to 42-52 (Vanhamme *et al.*, 2000), but only one is active at any one time. The BESs are large (~50kb) polycistronic transcription units that appear to be exclusively telomeric and contain not only the *VSG* gene but also a number of other genes termed ES associated genes (*ESAGs*). So far, 12 ESAGs have been identified (reviewed in Pays *et al.*, 2001), but, their presence or absence, order and copy number varies between individual BES (Fig. 1.3). Only *ESAGs* 6 and 7 appear to be present in all BES sequenced to date (Berriman *et al.*, 2002) and, as a result, are the only ESAGs to be constitutively expressed in the bloodstream. Although 12 ESAGs have been identified, the functions of only three of them have been defined to date. ESAGs 6 and 7 have been shown to encode a heterodimeric membrane-bound transferrin receptor (Ligtenberg *et al.*, 1994; Salmon *et al.*, 1994; Steverding *et al.*, 1994), and ESAG 4 has been shown to encode an adenylate cyclase (Alexandre *et al.*, 1996). One gene, found in *T. b. rhodesiense*, was named the serum resistance-associated (*SRA*) gene, before it was shown to that it could also be present as an *ESAG* (De Greef and Hamers, 1994). SRA confers resistance to human serum, encodes a VSG-related protein and is present in some, but not all, BES (Xong *et al.*, 1998). More recently, SRA has been shown to be a lysosomal protein that confers resistance by interacting with the serum protein apolipoprotein L-1 (Vanhamme *et al.*, 2003).

The region in the BES between the *ESAGs* and the *VSG* contains a large repetitive region, containing many degenerate copies of a single sequence, known as the 70-bp repeats (Liu *et al.*, 1983). The 70-bp repeats appear to be exclusively associated with the *VSGs*, both in the ES and in silent copies elsewhere, and often demarcate the boundary of recombination (Liu *et al.*, 1983; Matthews *et al.*, 1990). It has been suggested that these are a sequence specific element that promotes VSG switching (Barry, 1997), but other work in which the 70-bp repeats were deleted from an active BES suggest that they are not essential for recombination (McCulloch *et al.*, 1997). The BES promoter is well characterised upstream of the *ESAG* genes and directs RNA polymerase I-mediated transcription (Gunzl *et al.*, 2003). Another repeat region, known as the 50-bp repeats, is found upstream of the promoter of all known BES. This repeat region can extend for 40 – 50 kb and is thought to function as a barrier, separating the upstream sequences from the regulatory forces of the BES (Sheader *et al.*, 2003).

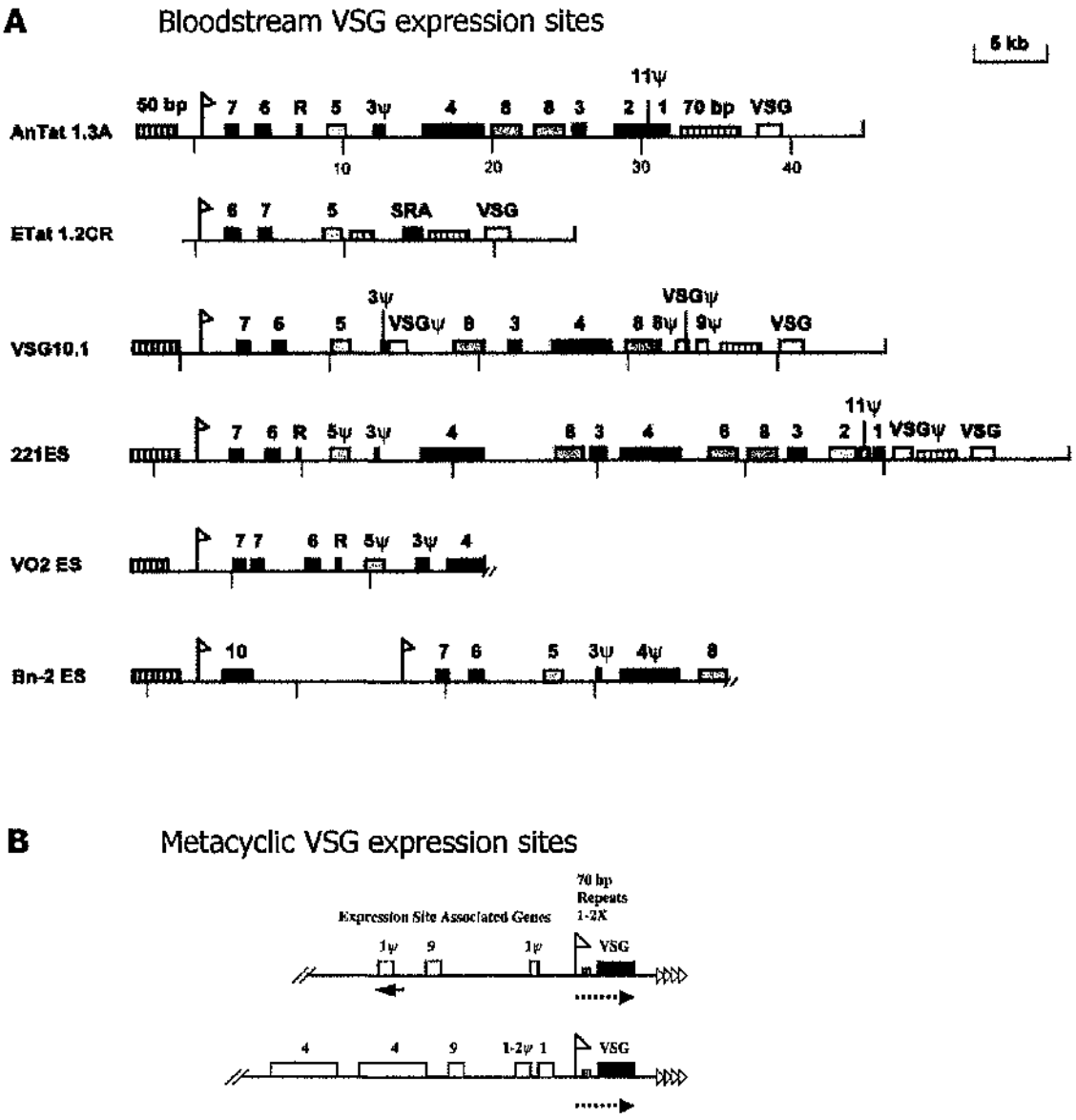


Figure 1.3: The expression sites of *T. brucei*. Schematic representations of the expression sites found in *T. brucei*. A: The BESs shown are AnTat 1.3A, Etat 1.2CR, VSG10.1, 221 ES, VO2 ES and Bn-2 ES. The ES promoters are indicated by a white flag, the 50-bp repeats by thickly striped boxes, the 70-bp repeats are thinly striped boxes, the *ESAGs* and the pseudo-*ESAGs* (ψ) by black and grey boxes respectively, and white boxes represent the *VSGs*. B: Two MESs are also shown. The ES promoters are indicated by a white flag, the 70-bp repeats by striped boxes, the *ESAGs* and the pseudo-*ESAGs* (ψ) by white boxes and black boxes represent the *VSGs*. The BES figure was taken from Berriman *et al* (2002) and the MES figure from Rudenko (2000).

MESs, in contrast to the BESs, are monocistronic transcription units only a few kb in length. It is predicted that as many as 27 MESs may be present in the genome (Turner *et al.*, 1988), and these also appear to have a telomeric location (Alarcon *et al.*, 1994). The *VSG* within the MES is, like in the BES, located most proximal to the telomere. The MES promoter, which is distinct in sequence from the BES promoter (Alarcon *et al.*, 1994; Ginger *et al.*, 2002), is situated approximately 2 kb upstream of the *VSG*. The region between the promoter and the *VSG* does not contain any *ESAGs*, but does contain 70-bp repeats, although not to the same extent as the BES, as only 1 or 2 repeats are normally present (Matthews *et al.*, 1990). The presence of numerous *ESAGs* and *ESAG* pseudogenes present upstream of the MESs has led to suggestions that they have arisen from the BES (Graham *et al.*, 1999; Rudenko, 2000).

Transcription of the ESs is highly unusual in that it is mediated by RNA polymerase I, whereas transcription of other protein coding genes is mediated by RNA polymerase II (Gunzl *et al.*, 2003). Also, the expression of the *VSGs* appears to be regulated at the transcriptional level (Vanhamme *et al.*, 2000), which is in contrast to other protein coding genes, such as the glycolytic enzymes, whose expression is regulated in a posttranscriptional manner (Vanhamme and Pays, 1995).

In bloodstream form *T. brucei*, the ability to regulate transcription so that only one BES is active at any one time is almost certainly linked to the discovery of a sub-nuclear entity known as the expression site body (ESB) (Navarro and Gull, 2001). The ESB is distinct from the nucleolus and may be a discrete proteinaceous structure (Chaves *et al.*, 1998; Navarro and Gull, 2001). The ESB also appears to contain RNA polymerase I and only the active BES, suggesting that it has a role in antigenic variation. This conclusion is strengthened by observations that the ESB has only been shown to be present in bloodstream form cells (Navarro and Gull, 2001). Exactly how the ESB contributes to *VSG* transcriptional control is not known. However, the localisation of only one BES within the ESB may allow it to be fully transcribed, with RNA elongation inhibited at all other BES (Vanhamme *et al.*, 2001a). This suggestion has been made in the light of observations that the initiation of transcription occurs at the majority of BES (Ansorge *et al.*, 1999; Vanhamme *et al.*, 2000) in bloodstream form cells.

Another process that may help explain the full expression of only one BES is the discovery of a DNA modification, called base J (Gommers-Ampt *et al.*, 1991; Gommers-Ampt *et al.*, 1993; van Leeuwen *et al.*, 1996). β-D-glucosyl-hydroxymethyluracil (base J) replaces ~0.5-1% of thymine in the *T. brucei* genome and is conserved within the Kinetoplastida and in Euglena (Dooijes *et al.*, 2000; van

Leeuwen *et al.*, 1998). The function of base J has not yet been defined, but it has been localised to telomeric repeats, other repeated sequences elsewhere in the *T. brucei* genome (van Leeuwen *et al.*, 2000), and in *VSG* and other BES sequences exclusively within the inactive BESs (van Leeuwen *et al.*, 1997). The absence of base J from silent *VSG* arrays and from the active BES have led to the suggestion of a role in BES silencing (van Leeuwen *et al.*, 1996). It is possible base J is a direct block to transcription and that the ESB is involved in removing the modified base from an inactive BES during switching, resulting in its activation. Alternatively, base J might be a second layer of transcriptional repression over the absence of an ESB directing transcription of the inactive BES. The fact that base J is not limited to *T. brucei*, or to VSG-related sequences within *T. brucei*, may mean that base J has a function outwith BES transcriptional control. It is known that at least one kinetoplastid protein, JBP1, binds base J, and elucidating the function of this may clarify the contribution of base J to BES control (Cross *et al.*, 1999; Cross *et al.*, 2002).

1.3.3 Mechanisms of VSG switching in *T. brucei*

The mechanisms employed to mediate VSG switching in bloodstream form *T. brucei* fall into two categories: transcriptional or recombinational. What is known about these mechanisms, and their relative significance, is discussed below.

1.3.3.1 Transcriptional (*in situ*) switching

Transcriptional switching appears to involve the co-ordinated inactivation of full transcription of the active BES and switching on full transcription in a previously inactive BES (Fig. 1.4). This process differs from all other VSG switching mechanisms, as it appears not to involve the movement of a VSG by recombination. Although DNA rearrangements have been reported to be associated with *in situ* switching (Gottesdiener *et al.*, 1992; Navarro and Cross, 1996), the fact that no specific DNA rearrangement was determined would suggest that it is not a requirement of transcriptional switching.

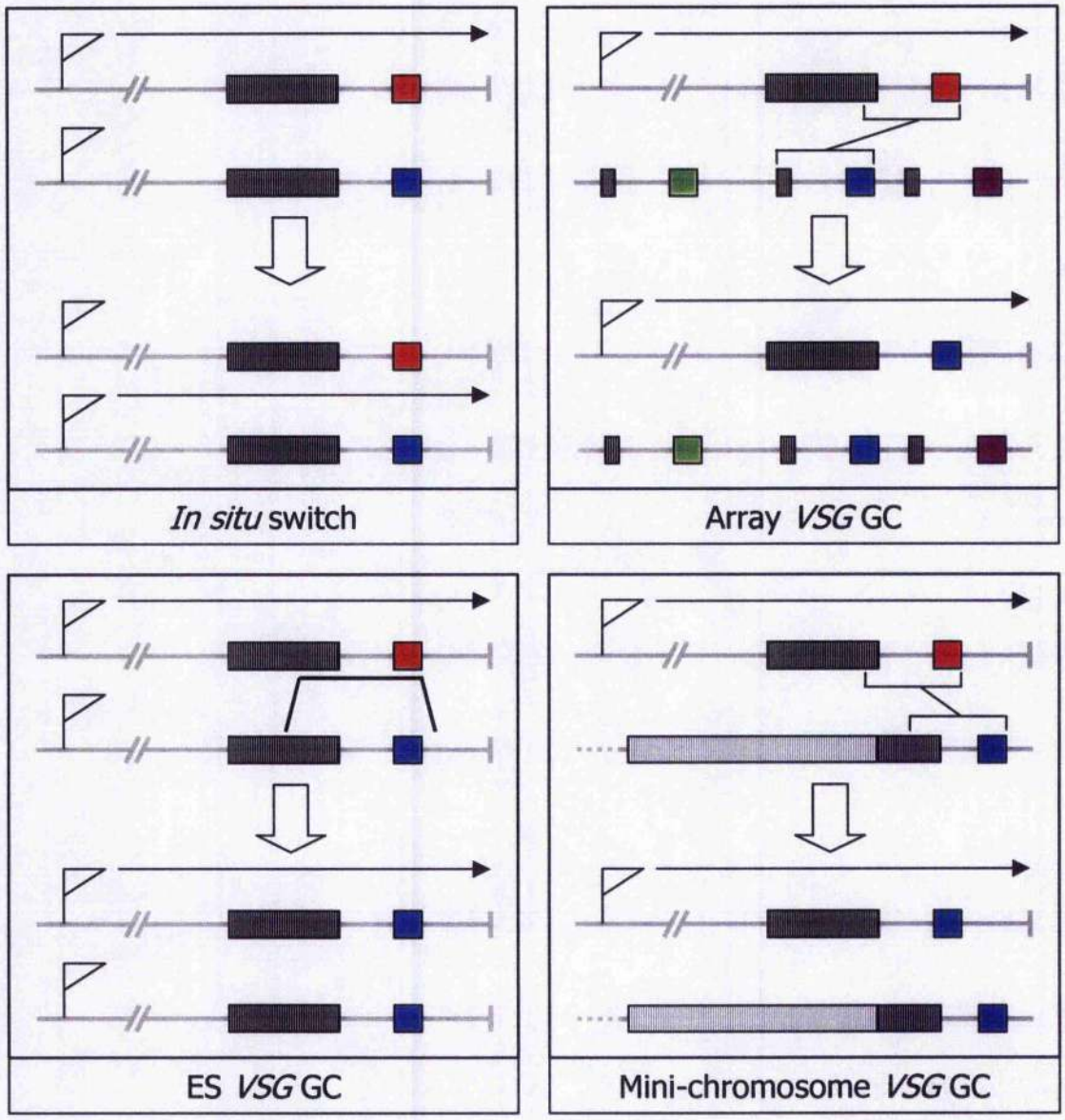


Figure 1.4: The VSG switching mechanisms of *T. brucei*. A schematic representation of the mechanisms of VSG switching. See text for an explanation of the switching mechanisms. Horizontal grey line: expression site and the *VSG* array chromosome. Vertical grey line: the end of the telomere. Triangle: the expression site promoter. Black arrow: transcription of the expression site. Black and white striped box: the 70-bp repeats. Grey and white striped box: the 177-bp repeats. Coloured squares: distinct *VSGs*. Black lines: the extent of sequence copied into the active expression site. GC: gene conversion.

It is still unknown how in situ transcriptional switching occurs, or might be regulated. A degree of regulation of the reaction, or 'cross-talk' between the BESs, has been suggested by experiments which attempt to select for *T. brucei* cells expressing two BESs (Chaves *et al.*, 1999). Full transcription of two BESs appears impossible, and instead this work uncovered an unsteady state, in which VSG expression appeared to rapidly fluctuate between the two BESs (Chaves *et al.*, 1999), possibly indicating an intermediate that occurs during *in situ* switching. The inability to isolate such putative intermediates between three BES (Ulbert *et al.*, 2002a) suggests that this 'cross-talk' can only extend to two BESs, the active site and the site about to be activated. This work appears to argue against there being any form of random activation or inactivation, as initially postulated by telomeric silencing models (Horn and Cross, 1995; Rudenko *et al.*, 1995). However, it is not clear how this relates to the function of the ESB, or base J. Moreover, the *trans* or *cis* acting elements which mediate such cross-talk are unknown.

1.3.3.2 Recombinational switching

As the majority of *VSGs* are located outwith ESs, they must be copied or transposed into the active ES in order to be expressed. This must occur by recombination, and several distinct mechanisms of VSG switching that involve recombination have been described.

Duplicative transposition is a gene conversion reaction, which is defined as the non-reciprocal transfer of genetic information from one DNA molecule to its homologue. Duplicative transposition involves the replacement of the *VSG* within the active expression site with a copy made from a silent *VSG*. This can result in the activation of a silent *VSG* from an array location (array VSG GC; Fig. 1.4), from another ES (ES VSG GC; Fig. 1.4), or from a mini-chromosomal location (MC VSG GC; Fig. 1.4). A silent *VSG* can also be recombined into the active ES by using sequences upstream of the 70-bp repeats for homology in a process known as ES gene conversion (ES GC; Fig. 1.5). The last type of gene conversion reaction involves the replacement of sequence in the active ES from a telomeric location (telomere conversion; Fig. 1.5).

The length of sequence transposed during gene conversion is variable, compatible with the reaction being driven by homologous recombination (Rudenko *et al.*, 1998). To achieve the successful recombination of the full VSG, homologous recombination utilises sites of homology upstream and downstream of the *VSG* to direct homologous

recombination. During *VSG* gene conversion (*VSG* GC) the 70-bp repeats normally mark the site of 5′ homology for transfer of the *VSG* 'cassette' (Liu *et al.*, 1983; Matthews *et al.*, 1990). However, deletion or inversion of the 70-bp repeats was shown to have no effect on the level of *VSG* GC (McCulloch *et al.*, 1997). This suggests that some gene conversions can use upstream sequences, probably the *ESAGs*, rather than the 70-bp repeats (Donelson *et al.*, 1983; Kooter *et al.*, 1988; LaCount and Donelson, 2001; Pays *et al.*, 1985). The 3′ region of homology can take the form of the C-terminal coding region (Donelson *et al.*, 1983; Liu *et al.*, 1983; Liu *et al.*, 1985; Pays *et al.*, 1985) of the *VSG*, or the 3′ untranslated region (Timmers *et al.*, 1987).

The transfer of a *VSG* in to the active ES by a telomeric conversion mechanism can also use sites of homology further upstream of the *VSG* (Kooter *et al.*, 1988; Lee and Van der Ploeg, 1987), and has also been shown to extend as far downstream as the telomere repeats (de Lange *et al.*, 1983; Scholler *et al.*, 1989). This mechanism of activation, however is limited to the silent *VSGs* located on the telomeres of mini-chromosomes or in other BESs.

Reciprocal telomere exchange has also been observed during VSG switching (Rudenko *et al.*, 1996). This switch mechanism involves the exchange of telomeres with no loss or gain of sequence (Fig. 1.5). However, as is the case for telomere conversion, it is limited to the activation of silent *VSGs* located on the telomeres of mini-chromosomes or in other BESs, and it is most likely a less frequent process.

One other mechanism of recombinational switching is mosaic gene formation, which involves the creation of a new *VSG* from fragments of existing genes or pseudogenes (Fig. 1.5). Pseudogenes are fragments of *VSGs* or non-functional *VSGs*, which can only be activated through segmental gene conversion into the active ES, creating a composite gene. This event is thought to occur late in infection to prolong chronic infections after exhaustion of the intact *VSG* repertoire, and is therefore considered to be a rare event (Kamper and Barbet, 1992; Pays, 1989; Thon *et al.*, 1989). Recent analysis of the genome has shown there to be a larger number of pseudogenes than first thought (L. Marcello and J.D. Barry, pers. comm.). As a result, it is likely that mosaic gene formation may account for a more significant proportion of switching events than previously estimated, possibly by increasing the *VSG* repertoire beyond the physical number of *VSGs* in the genome.

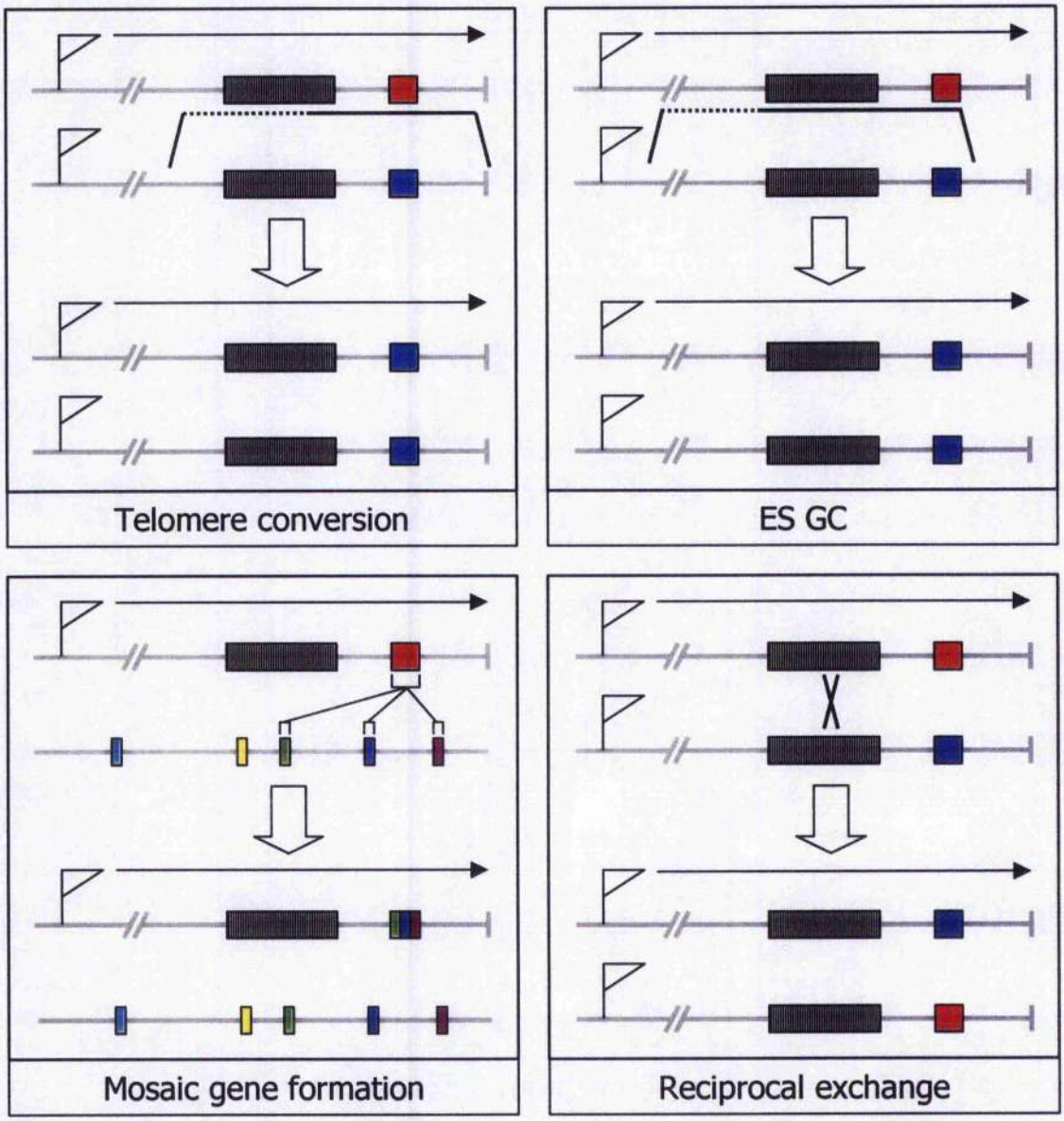


Figure 1.5: The VSG switching mechanisms of *T. brucei* (cont.). A schematic representation of the mechanisms of VSG switching. See text for an explanation of the switching mechanisms. Horizontal grey line: expression site and the *VSG* array chromosome. Vertical grey line: the end of the telomere. Triangle: the expression site promoter. Black arrow: transcription of the expression site. Black and white striped box: the 70-bp repeats. Coloured squares: distinct *VSGs*. Coloured rectangles: *VSG* pseudogenes. Black lines: the extent of sequence copied into the active expression site, which can be variable (dotted line).

1.3.3.3 The relative use of switching mechanisms

Monomorphic *T. brucei* cells, which have been used in the majority of switching studies to date, appear to predominantly utilise *in situ* switching, at least early in infection (Liu *et al.*, 1985). This is perhaps due to rapid serial passaging, which may require a reduced repertoire of switching compared to chronic cattle infections. However, later in infection the use of recombination to activate silent *VSGs* plays a larger role (Timmers *et al.*, 1987). There is no evidence that trypanosomes can regulate the relative use of *in situ* and recombinational switching, meaning that this is most likely explained by a predominance of *in situ* switching and an uncovering of recombinational switching as antibodies are generated against VSG species in the BES. In contrast, *VSG* gene conversion is the switching mechanism that predominates early in infections with pleomorphic cell lines (Robinson *et al.*, 1999; L. Morrison, PhD Thesis). Monomorphic trypanosomes may have lost or down-regulated the function of a specific factor that interacts with the homologous recombination machinery resulting in its ability to regulate, initiate or carry out *VSG* gene conversion. If true, monomorphic cell lines may rely on general levels of recombination for *VSG* switching, and therefore rely upon the general recombination machinery.

A hierarchy of the activation of *VSGs* during an infection has also been described, with some genes prone to appear early in infection and some late (Robinson *et al.*, 1999). Unpublished work by L. Morrison (PhD Thesis), showed that *VSGs* contained within telomeric locations are preferentially activated; this is possibly due to the close proximity of the telomeres within the nucleus and the level of homology between telomeres (Freitas-Junior *et al.*, 2000). It could also be that all the VSGs present in telomeres are full length functional. Within the group of telomeric *VSGs*, those located on mini chromosomes appear to be activated with slight preference to those on the larger chromosomes, probably as a result of their larger copy number. Finally, in the hierarchy, *VSG* activations from array locations (Robinson *et al.*, 1999; L. Morrison, unpublished) appeared to be the least frequent. Neither of these studies identified mosaic *VSGs* during switching analysis, suggesting that these remain the latest events in the hierarchy.

1.4 DNA double strand break repair

DNA double strand breaks (DSBs) can be caused by a plethora of potentially mutagenic agents. These include free radicals, ionising radiation, chemicals, transposons, blocks to DNA replication, and specifically-targeted cellular enzymes. Unrepaired, DSBs could result in chromosomal fragmentation, translocations and deletions, with long-term consequences including death and cancer (in multicellular organisms) (Jackson, 2001). DSB repair is therefore an essential process that acts to maintain genomic integrity. The mechanisms of double strand break repair can be divided into two categories: homologous recombination (reviewed in Krogh and Symington, 2004; Paques and Haber, 1999; Symington, 2002) and non-homologous end joining (NHEJ) (reviewed in Lieber *et al.*, 2003; Weterings and van Gent, 2004). Homologous recombination repairs DSBs through DNA sequence homology and generally repairs DNA accurately. NHEJ, on the other hand, ligates the ends of a DSB, often resulting in sequence changes. Both mechanisms of DSB repair are discussed below.

More recently, it has become clear that potentially the main role of homologous recombination is the repair of stalled replication forks. Stalled replication forks can arise due to DNA damage in replication substrates, inhibition of the replication machinery, or the blocking of the replication machinery by DNA-protein complexes. Stalled replication forks represent a source of genomic instability of potentially lethal consequences, and can lead to the creation of DSBs (Michel *et al.*, 2004). The interplay between replication and recombination is still being established, and the mechanism of how the homologous recombination machinery is able to restart stalled replication forks is beginning to be uncovered. Bacteria appear to have at least two pathways of replication fork restart. The GAP repair mechanism restarts forks that have stalled after encountering a DNA lesion and forks that stall after encountering a DNA nick are restarted by the DSB repair mechanism, in both cases recombination is involved (Cox *et al.*, 2000). In eukaryotes the replication fork restart is much less clear, however, recombination is still proposed to play a major role (Krogh and Symington, 2004).

1.5 Non Homologous End Joining

NHEJ (reviewed in Lieber *et al.*, 2003; Weterings and van Gent, 2004) is the process of ligation of DNA ends formed as a result of a DSB (Fig. 1.6a), often involving very

small regions of sequence homology (2 to 4 bp). In mammalian cells, NHEJ begins with the assembly of the DNA-dependent kinase (DNA-PK) at the DSB. The Ku heterodimer, consisting of Ku70 and Ku80, is the first to bind to the ends of the DSB, either tethering them (Cary *et al.*, 1997), or providing end protection (Mimori and Hardin, 1986). The Ku heterodimer then recruits the catalytic subunit of DNA-PK (DNA-PK_{cs}), which is activated by Ku (Hammarsten and Chu, 1998; West *et al.*, 1998). DNA-PK has been shown to phosphorylate a number of proteins, including p53, Ku, Xrcc4 and DNA-PK_{cs} itself (Smith and Jackson, 1999). How exactly DNA-PK is involved in the process of NHEJ is currently a matter of conjecture. Next, the DNA ligase IV/Xrcc4 complex is recruited to the DSB forming a tetrameric structure (Sibanda *et al.*, 2001), in which ligation of the DNA ends occurs.

The same core proteins appear to be involved in yeast NHEJ, although no DNA-PK_{cs} has yet been identified (Pastwa and Blasiak, 2003). Whether another kinase assumes this function, or if the reaction proceeds in its absence, is not known. Interestingly, the proteins of the MRX complex (see Section 1.6.2) are essential for NHEJ, as disruption of the individual genes results in a defect in NHEJ (Moore and Haber, 1996), where as the equivalent complex, MRN, is not required in vertebrates (Yamaguchi-Iwai *et al.*, 1999).

Very recently, it has been suggested that NHEJ might be conserved in bacteria also (Weller *et al.*, 2002; Wilson *et al.*, 2003). Several bacterial Ku homologues have been identified and have been shown to recruit DNA ligase to DNA ends and promote ligation (Weller *et al.*, 2002). The single copy Ku found in bacteria was shown to be homologous to the central regions of the eukaryotic Ku70 and Ku80 genes, suggesting that a gene duplication occurred to result in the two copies found in eukaryotes (Wilson *et al.*, 2003). Also, as is the case in eukaryotes, bacteria were shown to require Ku for fully functioning NHEJ (Wilson *et al.*, 2003).

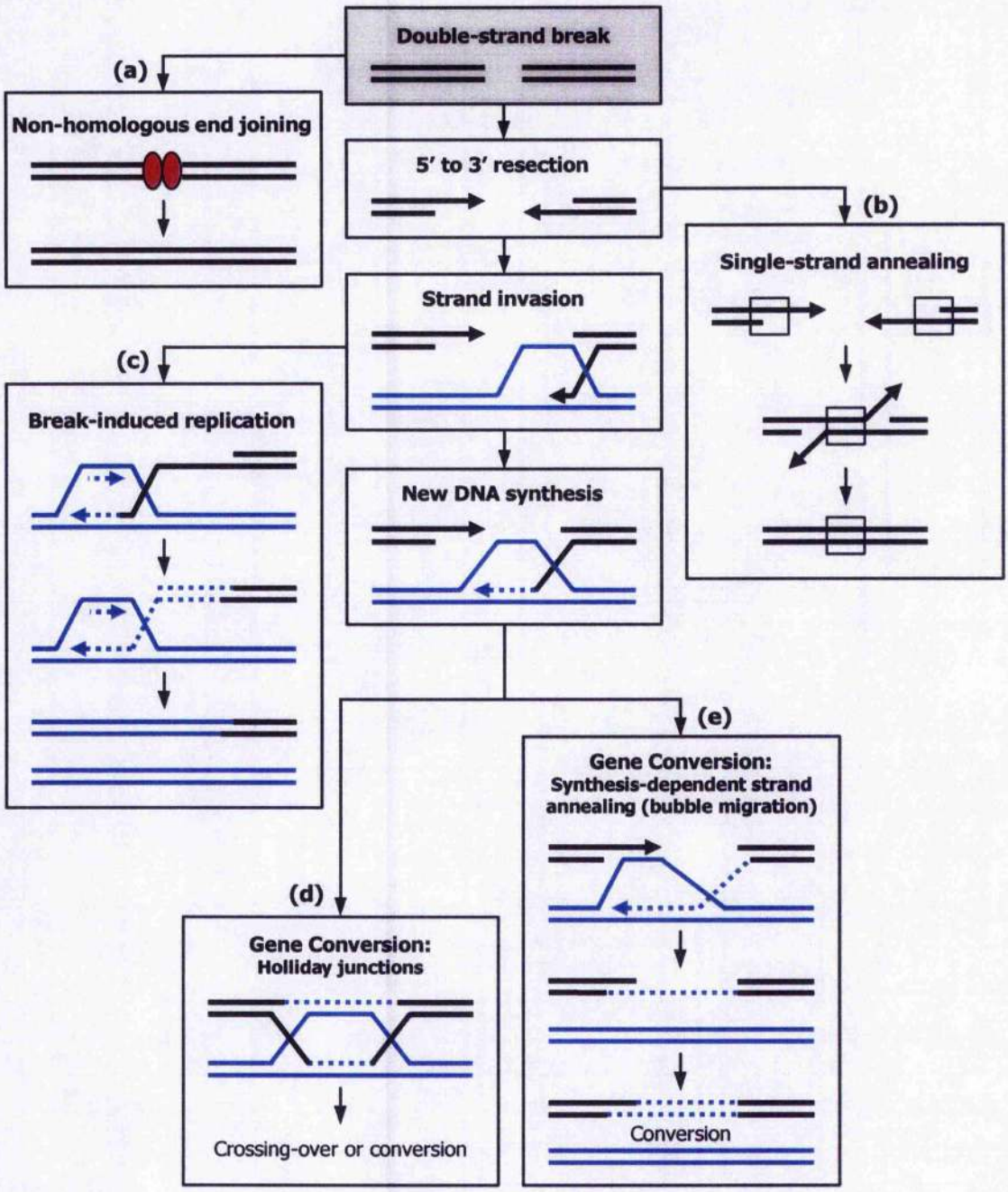


Figure 1.6: Pathways of double strand break repair. Black lines: duplex DNA with a DSB. Blue lines: intact duplex DNA. Dashed lines: newly synthesised DNA. Red circles: the NHEJ proteins. Black box: regions of homology. See text for details.

1.6 Homologous recombination

There are several processes of DSB repair that occur by homologous recombination and are conserved from bacteria to humans (Cromie *et al.*, 2001; Fig. 1.6). The underlying principles appear to be similar, however, in that each utilises broadly the same proteins, and follows essentially the same catalytic steps (see Krogh and Symington, 2004; Paques and Haber, 1999; Symington, 2002) for a comprehensive review). The mechanism of homologous recombination can broadly be split into three distinct catalytic steps: presynapsis, synapsis and postsynapsis (Hamatake *et al.*, 1989; Fig.1.7). The first stage, presynapsis, involves the processing of the ends of the DSB to provide a substrate for homologous recombination. This is carried out by resection of the 5′ end of the DSB to leave a 3′ single strand overhang (Cao *et al.*, 1990). In synapsis, the 3′ overhang invades homologous DNA and creates recombination intermediates; this step is often known as strand exchange. Termination of homologous recombination, or postsynapsis, can be mediated through a number of different pathways (discussed below), initially depending on whether or not both ends of the DSB have invaded duplex DNA. Invasion of both ends of a DSB can result in the formation of four-stranded branched DNA structures known as Holliday junctions (Holliday, 1964), which must be resolved by specific enzymes. The concept of formation and resolution of Holliday junctions was developed in a model termed the DSB repair model (Szostak *et al.*, 1983). In an alternative model, the invading DNA strand can be expelled from the duplex DNA and re-anneal with the broken end DNA in a process termed synthesis-dependant stand-annealing (SDSA) (Nassif *et al.*, 1994). In both these models, gene conversion reactions occur, where there is nonreciprocal transfer of DNA sequence from the unbroken DNA to the molecule with the DSB.

Different forms of post-synaptic resolution can occur if the synaptic step proceeds in a different manner. Invasion of only one of the ends of the DSB can occur in a process known as break-induced replication (reviewed in Kraus *et al.*, 2001; Paques and Haber, 1999). In this case, a replication fork can be established that could potentially replicate the intact duplex DNA to the end of the chromosome. In another reaction, termed single stand annealing (reviewed in Paques and Haber, 1999; Symington, 2002), the broken ends do not invade intact DNA duplex, but rather homology is found around the DSB on the same molecule. Annealing then occurs and trimming, DNA synthesis and ligation complete the reaction.

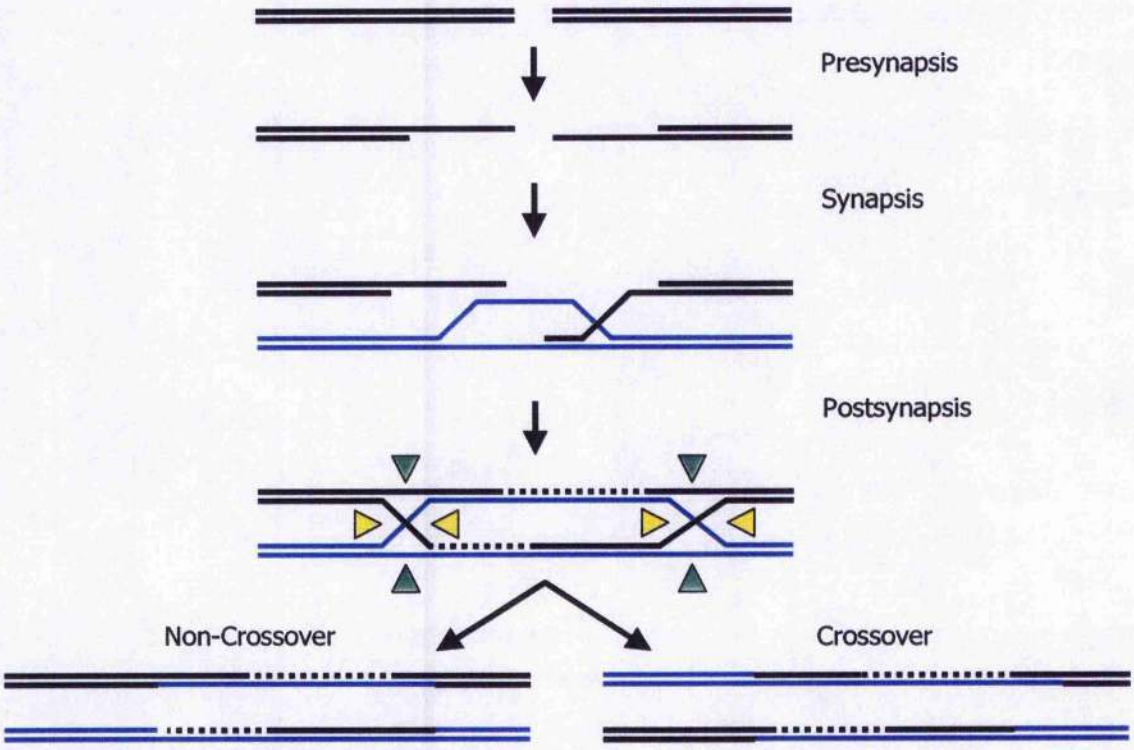


Figure 1.7: The double strand break repair model (Szostak *et al.*, 1983). A double strand break is present in a strand of duplex DNA (Black lines). During presynapsis the ends of the DSB are processed to provide 3' single strand overhangs. These subsequently invade intact duplex DNA (Blue lines) where they search for homology and create recombination intermediates in a process known as synapsis. During postsynapsis, DNA synthesis (Dashed lines) repairs any loss of sequence and the formation of Holliday junctions allows resolution of strand exchange intermediates. Resolution can either lead to a crossing over (both HJs being cut at points marked by different coloured arrowheads) or a non-cross over (both HJ being cut at points marked by the same coloured arrowheads).

1.6.1 Mechanisms of homologous recombination

Homologous recombination in eukaryotes utilises a large family of proteins initially identified in yeast and are known as the Rad52 epistasis group (reviewed in Symington, 2002). The members of this group are Rad50, Rad51, Rad52, Rad54, Rad55, Rad57, Rad59, Rdh54/Tid1, Mre11 and Xrs2, and these proteins are central to the different stages of recombination reactions discussed above. Some, but not all, of these yeast proteins are conserved from bacteria to mammals, and the discussion below will describe the function of these proteins and highlight similarities and differences in other eukaryotes and in bacteria.

1.6.2 Presynapsis

Presynapsis processes the ends of the DSB to provide a substrate for homologous recombination. This process is carried out in yeast by the Mre11/Rad50/Xrs2 (MRX) complex, which resects the 5′ ends of the DSB to leave 3′ single stranded DNA overhangs (Fig. 1.8). The primary evidence for this is suggested by the observation that 5′ to 3′ resection is attenuated in mutants of these genes (Ivanov *et al.*, 1994). The role of Xrs2 has been defined as an aid to the targeting of the MRX complex, providing a preference for single strand-double strand DNA junctions compared to single or double stranded DNA, and as a regulator of the exonuclease activity of Mre11 and Rad50 (Trujillo *et al.*, 2003). Mre11 and Rad50 form a sub-complex that exhibits a double stranded DNA 3′ - 5′ exonuclease activity (Trujillo and Sung, 2001). Since this activity is the opposite of that needed to create 3′ single strand DNA ends, the authors suggest that a helicase activity is required. Unwinding of the double stranded DNA at a DSB would result in the formation of 5′ and 3′ single strand DNA ends, and endonuclease cleavage of the 5′ strand by the MRX complex would leave the required overhang. Mre11 and Rad50 are conserved in all kingdoms of life, but the bacterial equivalents, SbcC and SbcD, do not have the same central functions in end-processing during homologous recombination (Connelly *et al.*, 1999). Xrs2 appears to be only present in eukaryotes, and is named Nbs1 in mammals (Symington, 2002). Analysis of the MR-Nbs1 complex in humans has suggested that resection of the DSB is carried out in a similar manner to yeast (Paull and Gellert, 1998). Mutants of Mre11, Rad50 and Xrs2 in yeast are all viable, whereas, Mre11, Rad50 and Nbs1 are all essential for viability in vertebrate cells (Symington, 2002). The MRX complex has many cellular functions besides end-processing (reviewed in Symington, 2002): it has also been shown to have roles in the maintenance of telomeres (Le *et al.*, 1999), in the creation and processing of DSBs during meiosis (Johzuka and Ogawa, 1995; Keeney, 2001) in signalling DNA damage (D'Amours and Jackson, 2002; Lisby *et al.*, 2004) and in NHEJ (van Gent *et al.*, 2001). DSB end processing is catalysed in bacteria by the RecBCD complex, which possesses 3′ - 5′ and 5′ - 3′ exonuclease activities (although the latter is much weaker), resulting in the preferential degradation of the 3′ end of a DSB. In addition, after the recognition of a chi site, the 3′ - 5′ nuclease activity of RecBCD is attenuated, resulting in the production of a 3′ single stranded DNA overhang to which RecA can bind (Anderson and Kowalczykowski, 1997; Kowalczykowski, 2000).

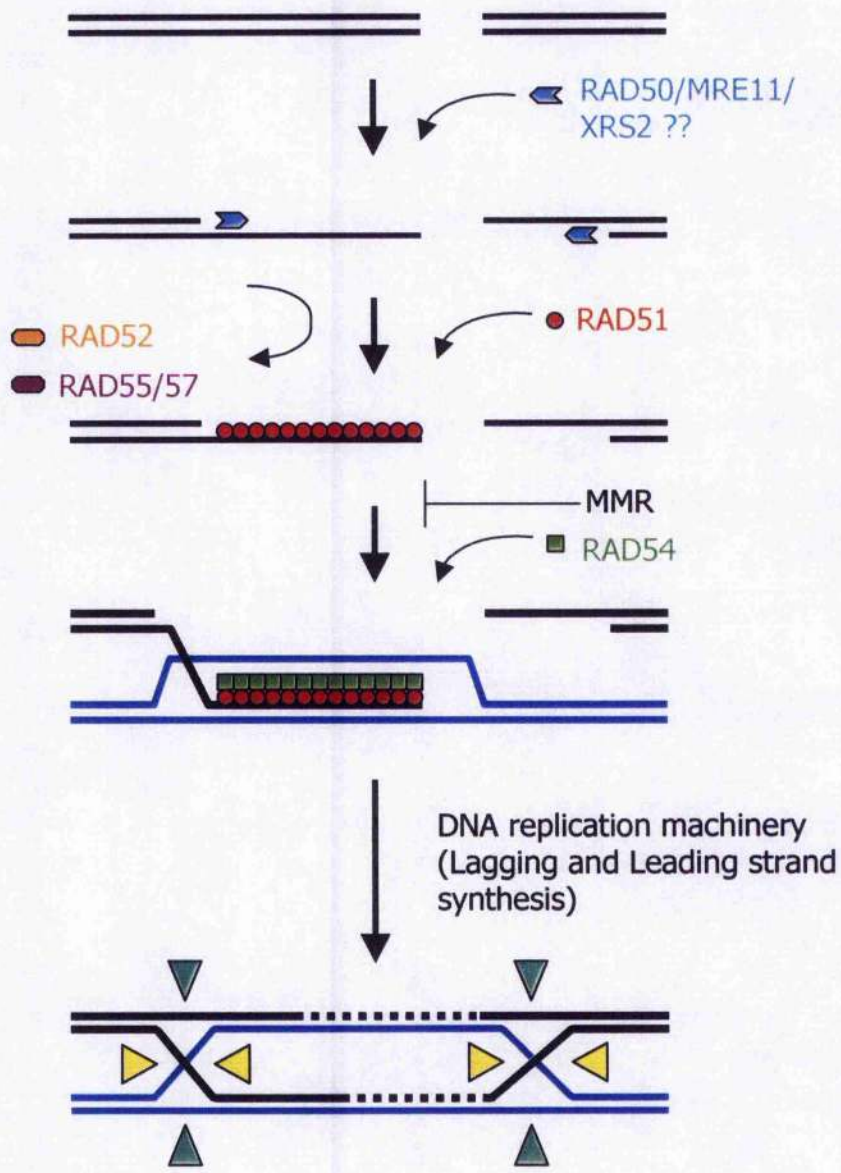


Figure 1.8: Some proteins involved in yeast homologous recombination. The black lines represent the duplex DNA that has suffered a DSB and the blue lines represent the intact duplex used as a template for the repair of the damaged strand. The dashed lines represent newly synthesised DNA and the coloured arrowheads mark the cut sites for the resolution on of the Holliday junctions. See text for further details of the proteins involved. MMR: Mismatch repair.

Before the process of synapsis can take place, a nucleoprotein filament must be formed on the single strand DNA to allow it to invade the homologous DNA. In eukaryotes the protein that this involves is Rad51 (Shinohara *et al.*, 1992), which is a structural and functional homologue of RecA in bacteria and RadA in archae (Brendel *et al.*, 1997). Rad51 is able to bind to both single strand DNA and double stranded DNA (Shinohara *et al.*, 1992), but Rad51 prefers tailed duplex DNA as a substrate for homologous recombination (Mazin *et al.*, 2000). The formation of a Rad51 nucleoprotein filaments results in a slight unwinding of the DNA, and 3 base pairs are bound to every Rad51 molecule (Sung and Robberson, 1995). The crystal structures of *S. cerevisiae* Rad51 and *E. coli* RecA nucleoprotein filaments have been solved (Conway *et al.*, 2004; Story *et al.*, 1992). Analysis of the pitch of Rad51 filaments shows then to have a pitch of 130 Å compared to 83 Å for the RecA nucleoprotein filament (Conway *et al.*, 2004). Rad51 is homologous to bacterial RecA, sharing 30% sequence similarity in the central conserved region (Shinohara *et al.*, 1992). Despite the low sequence conservation the nucleoprotein filaments formed by RecA (reviewed in Kowalczykowski, 2000) appear similar to Rad51 nucleoprotein filaments, although, they form with the opposite polarity (Shinohara *et al.*, 1993).

The loading of RecA onto single strand DNA requires DSB processing by the RecBCD complex, but no accessory factors appear to be needed. Why this is the case is as yet unknown. In contrast, and despite the fact that Rad51 can bind DNA on its own, the formation of a Rad51 nucleoprotein filament appears to be stimulated, or aided, in yeast by Rad52, Rad55, Rad57 (Fig. 1.8), and replication protein A (RPA). RPA is a heterotrimeric complex of RFA1, RFA2 and RFA3 and is thought to aid the loading of Rad51 onto single strand DNA by the removal of secondary structure (Sugiyama *et al.*, 1997; Sung and Robberson, 1995). The role of Rad52 appears to be to act as a mediator between Rad51 and RPA, aiding Rad51 to overcome the competition for single strand DNA posed by RPA (Sung, 1997a). It was originally though that Rad52 and Ku competed for DNA ends at a DSB, hence determining which pathway of DSB repair was taken (Van Dyck *et al.*, 1999). However, Rad52 interacts preferentially with single strand DNA, whereas Ku interacts preferentially with double stranded DNA ends, suggesting that this is not the case (Ristic *et al.*, 2003). Mutation of Rad52 results in the most severe recombination defects and DNA damage sensitivity of the entire Rad52 epistasis group (Paques and Haber, 1999). This is probably because Rad52 contributes to all forms of homologous recombination, including Rad51-dependent DSB repair, break-induced replication (see below) and single strand annealing (see below;

(Symington, 2002). Rad55 and Rad57 form a heterodimer which also acts to aid Rad51 nucleoprotein filament formation, in a manner similar to Rad52 function, since single mutants of all three genes fail to form Rad51 nucleoprotein filaments during meiosis (Gasior *et al.*, 1998). Mutation of Rad55 or Rad57 results in a radiation sensitivity and recombination defects which can be suppressed by the over-expression of Rad51 or Rad52 (Hays *et al.*, 1995; Johnson and Symington, 1995). In addition, the effects resulting from the loss of Rad55 and Rad57 can be partially suppressed by mutations in Rad51, meaning that Rad51 can act independently of these enzymes in some circumstances (Fortin and Symington, 2002).

In vertebrate cells, the number of proteins that contribute to Rad51 function is significantly more than in yeast. Five Rad51-related proteins have been identified (Rad51B, Rad51C, Rad51D, Xrcc2 and Xrcc3) which have been shown to form two, or perhaps three, discrete complexes (Liu *et al.*, 2002; Masson *et al.*, 2001b; Fig. 1.9). One complex is composed of Rad51C and Xrcc3 (the C3 complex), and the other (BCD2) contains Rad51B, Rad51C, Rad51D and Xrcc2 (Masson *et al.*, 2001b). It is proposed that the BCD2 complex results from the joining of two sub-complexes: Rad51B and Rad51C (BC) and Rad51D and Xrcc2 (D2) (Liu *et al.*, 2002). The same authors also propose that the formation of the C3 complex might arise from the interaction of Xrcc3 with Rad51C within the BCD3 complex. Clearly defined roles for the Rad51-related proteins and complexes are still emerging. However, mutants of any of these five genes, generated in chicken DT40 cells, were viable but exhibited chromosomal instability and recombinational repair defects, suggesting that they all have roles in repair and recombination (Takata *et al.*, 2001). On the other hand, mutants of Rad51B, Rad51D and Xrcc2 resulted in embryonic lethality in mice (Symington, 2002).

In humans, the formation of Rad51 nucleoprotein filaments, as in yeast, requires RPA (Baumann and West, 1998) and Rad52 (Benson *et al.*, 1998). It is proposed that this process is also aided by the Rad51B-Rad51C complex, in a process not dissimilar to that of the Rad55/57 heterodimer (Sigurdsson *et al.*, 2001). The Rad51C and Xrcc3 complex has been shown to bind Rad51 (Liu *et al.*, 2002). As yet no precise role has been suggested for Rad51B, Rad51D or Xrcc2 in DSB repair, although Rad51D has been shown to be involved in telomere maintenance (Tarsounas *et al.*, 2004b). Liu et al (2002) have proposed a model in which the 3′ single strand DNA tail is bound by the Rad51C-Xrcc3 sub-complex, either on its own or as part of the BCD2 complex. Subsequent binding of Xrcc3 to Rad51C, present as part of the BCD2 complex, would release Rad51B and the Rad51D-Xrcc2 sub-complex, leaving the Rad51C-Xrcc3

complex to recruit Rad51 to the DSB (Fig. 1.9). Indeed, it has been observed that Xrcc3 is recruited to the DSBs before, and independently of, Rad51 (Forget *et al.*, 2004). However, the Rad51B-Rad51C sub-complex has also been shown to possess Rad51 binding activity, although this appears to be weak (Lio *et al.*, 2003). It is possible that Rad51B binding to Rad51C partially obscures the Rad51 binding domain, and the release of Rad51B caused by Xrcc3 binding could then free the Rad51 binding site to allow Rad51C to interact with Rad51 with higher affinity. It seems likely that further work will add new dimensions to the roles of these Rad51-like proteins. Nevertheless, it seems unlikely that they are simply functional analogues of yeast Rad55/57, as other activities have been described at other stages of homologous recombination (see below).

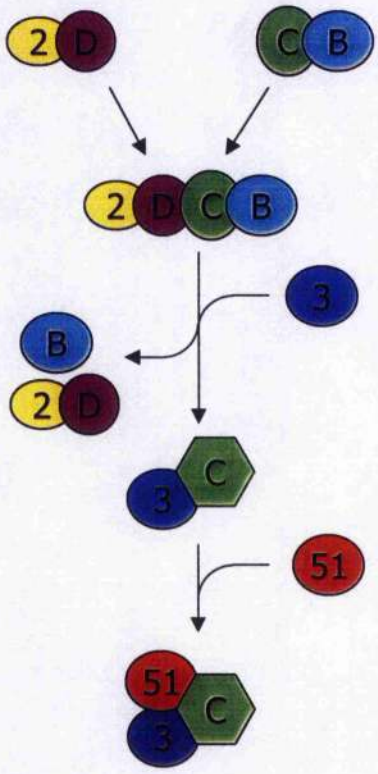


Figure 1.9: A model of the complexes of Rad51 and the Rad51-like proteins in humans. The diagram, which is a speculative model, takes into account all of the observed complexes discussed in the text. The proposers of this model suggest that a conformational change in Rad51C could account for the specificity of interactions. 51: Rad51. B: Rad51B. C: Rad51C. D Rad51D. 2: Xrcc2. 3: Xrcc3. See text for further details. Figure taken from Liu *et al* (2002).

Mammals also possess Rad52, which has been shown in humans to stimulate Rad51-dependant strand exchange (Benson *et al.*, 1998). However, in contrast to yeast, it is not so critical for repair and recombination in mammals (Yamaguchi-Iwai *et al.*, 1998). Another protein that is involved in the regulation of Rad51 function in mammals is Brca2 (Pellegrini and Venkitaraman, 2004). As yet no homologue of Brca2 has been identified in yeast, although, one was recently identified in *Ustilago maydis* (Kojic *et al.*, 2002) and many other eukaryotes (Lo *et al.*, 2003). Brca2 contains a number of 30 –

40 amino acid repeat sequences, known as the BRC repeats, which have been shown to bind Rad51 and be essential for resistance to DNA damage (Chen *et al.*, 1998). Brca2 has also been shown to be involved in the sequestering of Rad51 and the regulation of its DNA binding activity (Davies *et al.*, 2001). This has led to suggestions that Brca2 regulates the transfer of Rad51 on to single stranded DNA (Pellegrini and Venkitaraman, 2004), and that it may also be involved in the removal of Rad51 later on in DSB repair (Shivji and Venkitaraman, 2004).

1.6.3 Synapsis

Our understanding of how the Rad51 nucleoprotein filament acts to find homologous sequence, and transfer base pairing between DNA molecules is not clear. However, it is thought that the alignment of the Rad51-single strand DNA filament with homologous sequence, often within intact duplex DNA, occurs not by the invading strand sliding along the duplex, but through random collisions (Adzuma, 1998). The recognition of homology during this process has been proposed to occur by one of two mechanisms. The invading DNA could recognise homology in a locally unwound region. Alternatively, the DNA could form a triple helix structure in which the invading single strand interacts with one or both strands of the duplex DNA (reviewed in Rao *et al.*, 1995). Once a region of DNA sequence homology has been found, the intermediate structure is stabilised by the formation of base pairing between the single strand DNA and one of the stands of the duplex DNA. Elongation of the joint molecule can then take place by DNA synthesis and/or migration. In yeast and humans this can occur in either a 5′ to 3′, or a 3′ to 5′, direction with respect to the single strand DNA (Namsaraev and Berg, 1998; Namsaraev and Berg, 2000; Sung and Robberson, 1995). This migration, however, has been shown to be impeded by non-homologous sequence (Namsaraev and Berg, 1998; Namsaraev and Berg, 2000; Sung and Robberson, 1995).

The ability of Rad51 to form D-loops (the bubble of unwound DNA on the intact strand generated by the invasion of the Rad51 nucleoprotein filament; Fig. 1.6), and to search for homology, on its own is poor. The discovery of the *in vivo* and *in vitro* interactions between Rad51 and Rad54 may explain, in yeast and mammals at least, how this is overcome (Jiang *et al.*, 1996). Rad54 was shown to stimulate the rate of Rad51-dependant pairing between homologous single-strand and double-stranded DNA molecules (Petukhova *et al.*, 1998). Rad54 appears to result in an alteration in the DNA

conformation, possibly through strand separation or unwinding (Petukhova *et al.*, 1999b), thus aiding Rad51 to overcome structural impediments that would otherwise limit homologous DNA pairing (Fig. 1.8). The ability of Rad54 to promote this reaction appears dependent on its interaction with the Rad51 nucleoprotein filament (Van Komen *et al.*, 2000). RPA appears to be required at this step also, maximising the efficiency of the reaction, because binding of RPA to free single strand DNA prevents it from interfering with, and inhibiting, homologous pairing (Van Komen *et al.*, 2002). In humans, Rad54 appears to operate in a similar manner as it does in yeast (Sigurdsson *et al.*, 2002). The ability of Rad51 and Rad54 to promote strand exchange in humans has also been shown to be enhanced by the presence of chromatin (Alexiadis *et al.*, 2004) when compared to previous findings using naked DNA as a substrate, suggesting that it may act to aid Rad51 to overcome chromatin.

Recent analysis has suggested that mammalian Rad51C might also be involved in the strand transfer step of homologous recombination, utilising a strand melting and separation activity (Lio *et al.*, 2003). If correct, this suggests that Rad51C operates at two stages: an early stage where it aids Rad51 nucleoprotein filament formation and a later step in DNA pairing. This process would most likely be inefficient; however, the presence of Rad51 bound to single strand DNA would promote the ability of Rad54 to unwind the duplex DNA, therefore promoting the reaction.

As in eukaryotes, little is known regarding strand invasion and homology searching during DSB repair in bacteria. However, it appears that RecA is capable of carrying out these processes on its own, but is dependent on RecBCD loading it onto single strand DNA in a chi dependent manner (Anderson and Kowalczykowski, 1997; Kowalczykowski, 2000)

1.6.4 Postsynapsis

Postsynapsis is the least defined of the three steps of homologous recombination. This process must begin with the initiation of DNA synthesis from the 3′ OH of the invading DNA strand. It has been suggested that this is carried out by a replication fork similar to that used during replication initiated from replication origins, since gene conversion events at the mating type locus in *S. cerevisiae* were observed to require both leading and lagging strand DNA synthesis (Holmes and Haber, 1999).

In bacteria, the resolution of the Holliday junction molecules is well understood. RuvA, RuvB and RuvC are required for Holliday junction processing (reviewed in West, 1997). The RuvAB complex is responsible for branch migration after formation of the Holliday junction (West, 1996). The Holliday junction is targeted by RuvA, which then enables RuvB to form around the two opposite arms of the Holliday junction. The RuvC protein is the endonuclease that cleaves the Holliday junction (Connolly *et al.*, 1991), in a process that acts in conjunction with the RuvAB complex (West, 1996). The resolution of the Holliday junction involves the introduction of two nicks into the DNA of like polarity at preferred cleavage sites (Figs. 1.6d, 1.7 & 1.8).

In eukaryotes, Holliday junction resolution is much less clear. In yeast, it was originally thought that Mus81, which forms a complex with Eme1, acts in a similar manner to bacterial RuvC (Boddy *et al.*, 2001). However, more recent evidence suggests that this is not the case, proposing a role for Mus81-Eme1 in meiotic crossing over, but not gene conversion (Smith *et al.*, 2003). In humans, two putative Holliday junction resolution complexes with different specificities have been described. One of these contains Mus81 and may not truly be a Holliday junction resolvase (Constantinou *et al.*, 2002). More recently, Rad51B was shown to have Holliday junction binding activity, suggesting that it might have a role in junction resolution or branch migration (Yokoyama *et al.*, 2003). More recently still, it has been shown that Holliday junction resolution activity is reduced in CHO cell extracts carrying mutations of Rad51C or Xrcc3, suggesting that they also might have a role(s) in this process (Liu *et al.*, 2004). In the same study it was also suggested that Rad51C is involved in branch migration as cell extracts depleted of Rad51C showed reduced branch migration, a defect that was recovered when the extract was supplemented with any of the Rad51C containing complexes (Liu *et al.*, 2004).

1.6.4.1 Synthesis Dependent Strand Annealing

Synthesis Dependent Strand Annealing (SDSA) is an alternative strategy for post-synaptic resolution of Rad51-catalysed recombination intermediates (Nassif *et al.*, 1994; Fig. 1.6e). This model was developed to explain the lack of crossing over events observed during mitotic gene conversion reactions, and is similar to the DSB repair model with two notable differences. One difference is that conservative DNA replication occurs, with all newly synthesised DNA ending up in the recipient molecule.

The second is that the mechanism does not require the use of Holliday junctions. In this form of gene conversion, both 3′ single strand DNA ends invade the duplex DNA and initiate DNA synthesis. The newly synthesised strands are then expelled from the duplex and annealing of the strands then rejoins the broken DNA molecule (Nassif *et al.*, 1994). One mechanism of SDSA that does allow crossing over was proposed by Ferguson and Holloman (1996). In their mechanism, they proposed that the migration of a single D-loop, and its subsequent annealing to the other end of the DSB, would result in the formation of a single Holliday junction that could either be resolved with or without a crossover event (Ferguson and Holloman, 1996).

1.6.4.2 Break Induced Replication

Break-induced replication (BIR), in contrast to gene conversion, requires invasion of only one end of the DSB (Fig. 1.6c). However, the initial steps of BIR and gene conversion are essentially the same. The 5′ end of the DSB needs to be resected to provide a 3′ single strand DNA overhang which then invades duplex and initiates replication. BIR, unusually, can still occur in the absence of Rad51, Rad54, Rad55 or Rad57, proteins that act in gene conversion. Mutation of these genes has been shown to almost abolish gene conversion, but BIR can still take place (Malkova *et al.*, 1996; Signon *et al.*, 2001). In contrast, mutation of rad50, rad59 and rdh54/tid1 (rad54B; a homologue of Rad54 that appears to act in the same manner, but in meiotic recombination) was shown not to adversely affect gene conversion except when coupled with a mutation of rad51 or rad54, which resulted in severe reductions of both gene conversion and BIR reactions (Signon *et al.*, 2001). These results suggest that Rad51-independent BIR requires Rad50 (presumably with the other components of the MRX complex: Mre11 and Xrs2), Rad59 and Rdh54/Tid1. How exactly BIR operates in the absence of Rad51 and its cofactors has yet to be determined. However, the requirement for Rad50 suggests that BIR may utilise the MRX complex to process the DSB to provide recombinogenic ends as it does in gene conversions. The only other requirement published to date is that of a 200 bp sequence located ~34kb from the DSB (Malkova *et al.*, 2001). How exactly this sequence aids BIR remains to be defined.

More recently, the process of Rad51-dependent BIR has been described. This reaction requires the same genes as gene conversion reactions (Rad52, Rad54, Rad55 and Rad57), and therefore has been suggested to have a common strand invasion

intermediate (Davis and Symington, 2004). This process has also been shown to be more efficient than the Rad51-independent BIR pathway and also to have no requirement for special facilitator sequences (Malkova *et al.*, 2005).

1.6.4.3 Single Strand Annealing

If a DSB occurs in between two repetitive sequences, repair of the break can result in a loss of DNA sequence in a process termed Single Strand Annealing (SSA) (reviewed in Paques and Haber, 1999). This mechanism arises due to the resection of both 5′ ends of the DSB, revealing homologous 3′ single strand DNA overhangs beyond the repetitive sequences. The exposure of two homologous regions allows them to be annealed to repair the break, but in the process lose one of the repetitive sequences and the sequence between them (Fig. 1.6b). Rad59 has been shown to be important for SSA between direct repeats, with the requirement for Rad59 increasing as the repeat length decreases (Symington, 2002). Rad59 has also been shown to bind DNA, preferentially single strand DNA, and be capable of annealing complementary single strand DNA (Petukhova *et al.*, 1999a). After the annealing of the repeats, excision of the non-homologous 3′ ends, DNA synthesis and ligation complete the reaction.

1.6.5 Meiotic recombination

Meiosis is a specialised form of cell division in some diploid organisms which involves a single round of DNA replication, followed by two subsequent cell divisions resulting in the generation of haploid cells (Page and Hawley, 2003). Meiotic recombination is an essential process of meiosis that stabilises the interactions between segregating chromosomes, generates diversity and promotes evolution. The mechanism of meiotic recombination is very similar to mitotic DSB repair, and appears to occur either by DSB repair or SDSA mechanisms (Smith and Nicolas, 1998). However, specific adaptations to this repair reaction are characteristic of meiosis (Fig. 1.10). During meiosis, DSBs are generated by the interactions of a number of proteins, with the topoisomerase-like Spo11 transesterase catalysing DSB formation (Keeney *et al.*, 1997). It is proposed that a pair of Spo11 monomers cleave DNA by a reversible transesterase reaction in which a tyrosine side chain on the protein attacks the phosphodiester backbone (Keeney *et al.*,

1997). Mutation of Spo11 in yeast has been shown to result in the absence of DSBs, a lack of Rad51-Dmc1 complex (see below), a lack of recombination and severe synaptonemal complex defects (Lichten, 2001). The generation of DSBs in addition to Spo11, requires a number of other proteins (Roeder, 1997). Whether or not the DSB generating complex assembles on the DNA as it is replicated is as yet still unknown. It is thought that the DSB is then resected, as it is in the DSB repair model, by the MRX complex (Keeney, 2001). In mice, humans, drosophila and plants homologues of yeast Spo11 have been identified (Hartung and Puchta, 2000; McKim and Hayashi-Hagihara, 1998; Romanienko and Camerini-Otero, 2000), and in mice Spo11 has also been shown to generate DSBs (Keeney *et al.*, 1999). This suggests that the process of generating DSBs to initiate meiosis is conserved from yeast to mammals.

Meiotic recombination also requires a meiosis-specific homologue of Rad51, named Dmc1 (Bishop *et al.*, 1992). Rad51 and Dmc1 proteins have high levels of sequence homology and both have distinct and overlapping roles in recombination (reviewed in (Masson and West, 2001). Dmc1 has been shown to form multiple nuclear complexes (known as foci) prior to meiotic chromosome synapsis (Bishop *et al.*, 1992), suggesting that it operates in a similar manner as Rad51 in DSB repair. The presence of Dmc1 has also been described in mammals, suggesting that the repair of the DSB is also conserved. Human Dmc1 has been shown to bind DNA in an octomeric ring (Passy *et al.*, 1999), interact with Rad51 (Masson *et al.*, 1999), and to be meiosis-specific (Gupta *et al.*, 2001). In contrast to the filaments formed by RecA and Rad51 on single stranded DNA, human Dmc1 binds single stranded DNA like beads on a string (Masson *et al.*, 1999).

Mutation of *Dmc1* in yeast has been shown to result in defective reciprocal recombination, an accumulation of DSBs and the failure to form normal synaptonemal complexes (SC), all of which caused the mutants to arrest late in meiotic prophase (Bishop, 1994). Recombination in meiosis is very distinct from mitosis in that it occurs preferentially between homologous chromosomes rather than sister chromatids. The mutation of another gene, *red1*, which is also required for SC formation, was shown to almost eliminate the production of inter-homologue recombination intermediates (Schwacha and Kleckner, 1997). In the same study, *red1 dmc1* double mutants were shown to have nearly abolished inter-homologue joint molecule formation, with Rad51 promoting joint molecules between sister chromatids. Together, these results suggest that the role of Dmc1 is to promote recombination between homologous chromosomes,

rather than sister chromatids, and that Red1 functions to block joint molecule formation in the absence of Dmc1.

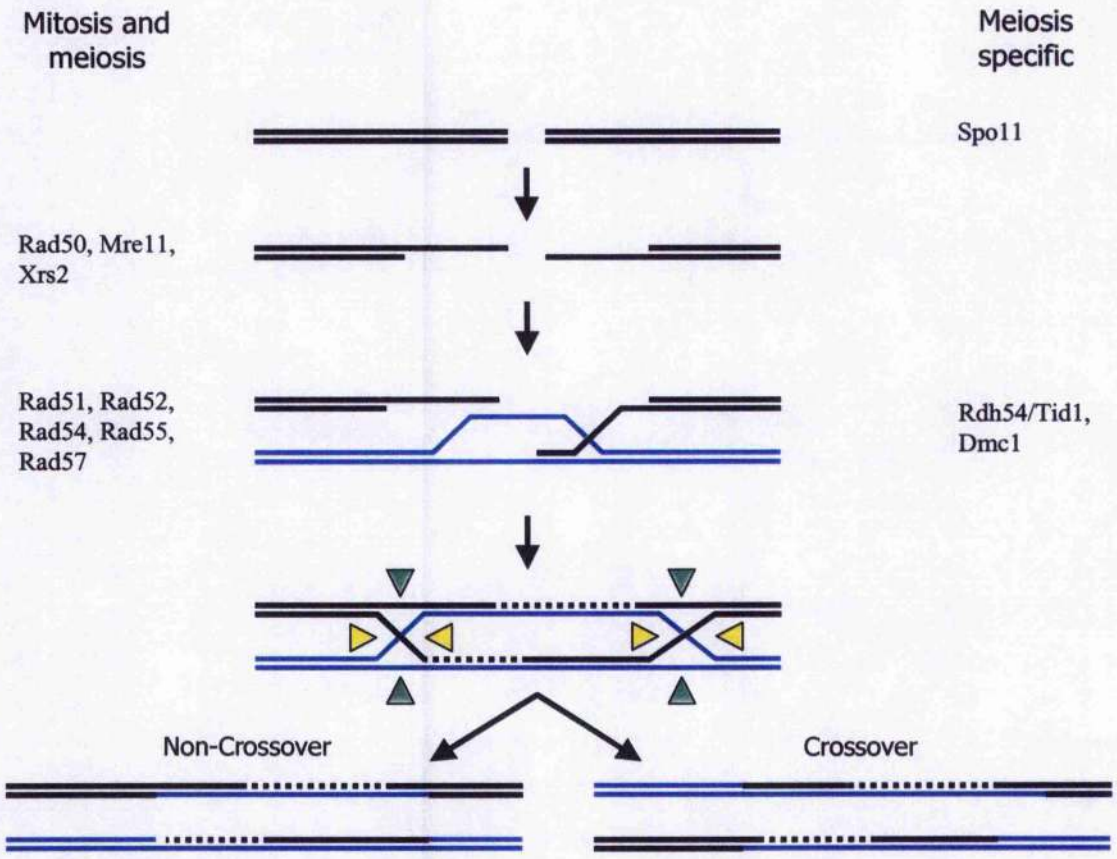


Figure 1.10: The proteins of mitotic and meiotic recombination. A schematic representation of DSB repair including a list of the proteins involved in mitotic and meiotic recombination, highlighting those that are meiosis-specific. The figure does not include all of the proteins involved in these processes, merely a subset of them that have been discussed during this introduction. Figure adapted from Smith and Nicolas (1998).

Rad51 and Dmc1 have been shown to co-localise during meiotic recombination (Bishop, 1994), and, in yeast, this appears to be promoted by a Rad54 homologue, Rdh54/Tid1 (Shinohara *et al.*, 2000). Although Rdh54/Tid1 has also been shown to operate in mitotic cells (Klein, 1997), it appears limited to aiding the process of recombination between homologous chromosomes, whereas Rad54 aids sister chromatid recombination (Arbel *et al.*, 1999). A homologue of yeast Rdh54 has also

been identified in mammals, named Rad54B, which has been shown to be involved in homologous recombination, although no direct interaction with Rad51 or Dmc1 was observed (Tanaka *et al.*, 2002). As yet, it has not been determined if mammalian Rad54B has a role in meiotic recombination.

The mismatch repair proteins, Mlh1, Mlh3, Msh4 and Msh5, have been shown to function during meiotic recombination (Roeder, 1997). Mutations in *msh4* and *msh5* result in a reduction of crossing over despite normal levels of gene conversion (Hollingsworth *et al.*, 1995; Novak *et al.*, 2001). Different roles for the Mlh and Msh proteins have been suggested by the observation that reduction in the levels of crossing over in *mlh1* and *mlh3* mutants is much less pronounced (Kleckner, 1996; Roeder, 1997). However, it appears that they may operate in the same pathway as the defects observed in an *mlh1 msh4* double mutant are no more severe than those of a *msh4* single mutant (Hunter and Borts, 1997). Although the exact roles of these proteins have not yet been defined, recent findings have begun to do so. Msh4 and Msh5 have been shown to recognise Holliday junctions and to form a sliding clamp that embraces adjacent homologous duplex DNA (Snowden *et al.*, 2004). Msh4 has also been shown to co-localise with Rad51 and Dmc1 (Neyton *et al.*, 2004). These results have led to a proposed function of Msh4 and Msh5 in the stabilisation and preservation of a DSB repair intermediate.

1.6.6 The role of Mismatch repair

During replication or mutagenesis by some chemicals, base pair mismatches can arise which could lead to phenotypic changes if left unrepaired. The mismatch repair (MMR) machinery, which is responsible for recognising such mismatches, has a vital role in the maintenance genomic integrity (reviewed in Schofield and Hsieh, 2003). In addition, it plays a role in homologous recombination (reviewed in Evans and Alani, 2000), since mismatches can arise during recombination between non-identical DNA sequences. In eukaryotes, the initial steps of MMR involve the binding of bacterial MutS homologues (Msh) to a mismatch and, in a reaction that requires ATP, recruitment of MutL homologues (Mlh, or Pms1 or 2). Together, these complexes co-ordinate the removal of the mispaired bases and DNA re-synthesis. Msh2 forms heterodimers with the other MutS homologues, with the resulting complexes functioning in different types of MMR. Msh2-Msh6 recognises single-base mismatches, whereas Msh2-Msh3 recognises 1-8bp

insertion/deletion loops. 4 Mlh proteins have been described, which form three different heterodimers: all these contain Mlh1, and one of either Mlh2, Mlh3 or Pms1 (Wang *et al.*, 1999). The distinctions between these heterodimers in the MMR reaction are yet to be elucidated

MMR also affects homologous recombination, inhibiting recombination between diverged sequences (Datta *et al.*, 1996; Elliott and Jasin, 2001). Exactly how MMR inhibits recombination is not fully understood, but it has been suggested that the inhibition could take place during the processes of strand assimilation and/or branch migration, with MMR preventing strand exchange or branch migration in diverged sequences (Datta *et al.*, 1996). Alternatively, MMR could result in the reversal of strand exchange intermediates, or destruction of intermediates with mismatches (Rayssiguier *et al.*, 1989).

1.7 DNA repair, recombination and antigenic variation in *T. brucei*

In an attempt to define the pathways of DNA repair and recombination in *T. brucei*, functional analysis has been carried out on a number of homologues of some of the repair and recombination proteins mentioned in previous sections. The following section summarises the findings to date, and discusses the implied influences of these proteins on the mechanism of antigenic variation in *T. brucei*.

The first gene that was shown to have a role in DNA repair and recombination in *T. brucei* was *RAD51*, a homologue of eukaryotic Rad51, bacterial RecA and archeal RadA (McCulloch and Barry, 1999). Disruption of the *RAD51* open reading frame, in both bloodstream form and procyclic stage cells, resulted in increased DNA damage sensitivity and a reduction in the efficiency of recombination, as measured by the rate of integration of transformed constructs. The same experiments also revealed unusual pathways of DNA recombination in *RAD51* mutants. These results, as expected, confirm that RAD51 indeed has a role in *T. brucei* repair and recombination. However, both processes were not abolished, suggesting that backup pathways must exist. Further analysis of the aberrant integrations taking place in the absence of *RAD51* suggested that at least one backup pathway of recombination utilises small regions of homology during recombination. These small regions of homology were generally 7 - 13 bp in length, and in most cases contained base mismatches or insertions/deletions (Conway *et al.*, 2002b). RAD51 has also been shown to be present in two other protozoan parasites;

L. major (McKean *et al.*, 2001) and *P. falciparum* (Bhattacharyya and Kumar, 2003). In both organisms, increased expression of RAD51 was observed in response to DNA damage, suggesting that as in trypanosomes, RAD51 is involved in DNA damage repair.

T. brucei also contains homologues of eukaryotic Ku70 and Ku80. Disruption of either gene in *T. brucei* had no effect on DNA damage sensitivity, with either methyl methanesulphonate or phleomycin as damaging agents (Conway *et al.*, 2002a). This suggests that either NHEJ is not present in *T. brucei*, or that homologous recombination dominates repair to such an extent that the NHEJ defect goes undetected. Determination of which of these possibilities is correct will require the generation of *RAD51* and *KU* double mutants. *T. brucei* KU70 and KU80, however, were found to be important in the maintenance of transcriptionally active telomeres, since telomere shortening was observed in either mutant. Surprisingly, transcriptionally inactive telomeres were less prone to length changes (Conway *et al.*, 2002a). This result suggested that the KU proteins are involved in a telomere maintenance pathway, a result confirmed by further analysis of KU80 in *T. brucei* (Janzen *et al.*, 2004). Despite the observed telomere maintenance role of the KU proteins, no *KU* mutant clone was shown to have lost all telomere sequence. In fact, a telomere length equilibrium was reached in *KU* mutants grown in culture for extended periods of time, suggesting that at least one alternative, and uncharacterised, pathway exists for the maintenance of telomeres in *T. brucei* (Conway *et al.*, 2002a).

Another protein shown to have important roles in DNA repair and recombination in *T. brucei* is Mre11. During the analysis of *MRE11* function in *T. brucei*, mutants were observed to show DNA damage sensitivity to phleomycin, but not MMS, and also altered efficiencies and integration patterns during recombination with transformed DNA constructs (Robinson *et al.*, 2002; Tan *et al.*, 2002). Although MRE11 in *T. brucei* appears to be involved in repair and recombination, further results point to a role that may be quite different from that defined in other organisms. Mutation of *MRE11* in *T. brucei* was shown to result in gross chromosomal rearrangements (GCRs) (Robinson *et al.*, 2002). These GCRs appeared to occur only through the deletion of chromosome-internal sequence as only a shortening of chromosome length was observed and the terminal fragments remained largely intact. This appears different from *MRE11* mutant GCRs seen in yeast, where translocations, telomere deletions and chromosomal-internal deletions occurred (Chen and Kolodner, 1999). Despite all of the megabase chromosomes appearing to be affected by GCRs in *T. brucei*, the smaller intermediate

chromosomes and the mini-chromosomes remained largely intact (Robinson *et al.*, 2002).

Five genes have been identified that, by sequence comparisons, are proposed to be involved in the MMR pathways of *T. brucei* (Bell *et al.*, 2004). Disruption of two of these genes, *MSH2* and *MLH1*, has shown that they are indeed involved in MMR, with loss of the genes resulting in increased rates of microsatellite variation and increased tolerance to *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (Bell *et al.*, 2004). Loss of *MSH2*, central to the 'recognition' MSH heterodimers of yeast, also resulted in an increased frequency of recombination between both homologous and diverged sequences (Bell and McCulloch, 2003). Therefore, as in yeast, the MMR pathways of *T. brucei* appear to act to prevent recombination between diverged sequences. However, a significant drop in the recombination efficiency of sequences showing just 1% divergence suggests that MMR does not account for all of this recombination inhibition.

The analysis of the mechanisms of recombination in other organisms allows us to predict how the process of recombinational switching might occur in trypanosomes. In theory the activation of array *VSG* genes requires a gene conversion reaction. It is most likely that this process would be carried out by SDSA (Barry, 1997; Borst *et al.*, 1996), a conservative replication method, as opposed to DSB repair which could result in translocations. Recombination of telomeric *VSG* into the active BES could proceed by gene conversion or BIR, as either would not lead to critical sequence loss since only telomeric sequences would be exchanged. Analysis of homologues of the proteins involved in these pathways in other organism could help start to define the mechanisms of recombinational switching in *T. brucei*.

To this end, disruption of *RAD51* was shown to result in a reduction in the rate of VSG switching (McCulloch and Barry, 1999), showing that recombinational switching is likely to involve one of the mechanisms described above. What was unexpected was that both *in situ* and recombinational mechanisms were affected. Why this is the case is not yet known as, *in situ* switches are thought not to involve recombination.

Unexpectedly, the disruption of *MRE11* resulted in no detectable defect in VSG switching (Robinson *et al.*, 2002). This is particularly surprising when it is considered that loss of MRE11 leads to a reduction in the ability to repair DSBs, the mechanism proposed to initiate VSG switching. This result could be explained by the initiation of DSB repair during VSG switching being carried out by a specialised endonuclease. Perhaps the form of DSB that is generated excludes any processing by MRE11. Alternatively, VSG switching could be initiated from a stalled replication fork in a

pathway that does not require MRE11. This is suggested by the observation that MRE11 mutants do not show increased sensitivity to MMS damage (Robinson *et al.*, 2002), which generates adducts on DNA that might be expected to induce replication fork stalling (Sedgwick, 2004), as opposed to phleomycin which directly creates DSBs (Giloni *et al.*, 1981). For either explanation to be correct, the proposed role that MRE11 has in signalling DNA repair must be absent in *T. brucei*.

Mutation of *KU* was also shown not to effect VSG switching (Conway *et al.*, 2002a), although this is perhaps not unexpected as NHEJ, a process that generates small levels of sequence alterations, is unlikely to be utilised in the life saving process of VSG switching. This reasoning would appear to rule out a transposition-based mode of *VSG* switching (Dujon, 1989), in keeping with the suggestion that RAD51 and homologous recombination drives VSG switching.

Despite influencing homologous recombination, MSH2 appears also not to influence VSG switching (Bell and McCulloch, 2003). There could be several explanations for this observation. The simplest explanation for this is that antigenic variation is driven largely by homologous recombination on perfectly matched sequences. If this is the case, the switching assay may not be sensitive enough to detect the slight increase in recombination efficiency observed in MMR-deficient *T. brucei* (Bell and McCulloch, 2003). However, we know that *VSG* gene conversion often utilises degenerate 70-bp repeats as part of the integration cassette. Alternatively, VSG switching may utilise a homologous recombination pathway exempt from regulation by MMR. This may be achieved by the exploitation of a novel homologous recombination pathway to carry out VSG switching. We already know that *T. brucei* utilises at least two pathways of homologous recombination, RAD51-dependent and RAD51-independent (Conway *et al.*, 2002b).

The analysis of VSG switching in the mutants described above have all been carried out in a derivative of Lister 427 bloodstream cells (McCulloch *et al.*, 1997; Rudenko *et al.*, 1996), a monomorphic *T. brucei* cell line that switches at only background levels. It is therefore difficult to say what effect, if any, these genes would have in pleomorphic strains with substantially higher VSG switching frequencies. Indeed, recent work analysing RAD51 in ILTat 1.2, a pleomorphic cell line, has suggested that in high switching strains its role in VSG switching may not be as central as it is in monomorphic cell lines (P. Burton, PhD thesis).

1.8 The aims of the project

Given that only one gene so far, *RAD51*, has been identified to have a role in antigenic variation, our picture of the genetic influences of this reaction is far from complete. In addition, *RAD51* mutant *T. brucei* display a DNA damage sensitivity, a recombination defect and a reduced rate of VSG switching, but since these processes do still occur, backup pathways of repair, recombination and VSG switching must exist. The broad aim of this thesis is the identification and analysis of other genes that may contribute to these processes.

During this introduction I have discussed the recombination mechanisms in mammals and yeast, and discussed the fact that the central enzyme, Rad51, is aided by a number of cofactors.

In light of this knowledge, the three specific aims of this project are as follows:

i. To identify *RAD51*-like genes present in *T. brucei*.

ii. To determine if any of the *RAD51*-like genes have roles in DNA repair and homologous recombination.

iii. To determine if any of the *RAD51*-like genes have roles in antigenic variation.

CHAPTER 2

MATERIALS AND METHODS

2.1 Reagent abbreviations

BSA	bovine serum albumin
CBSS	Carter's balanced salt solution 1x: 0.023 M HEPES, 0.12 M NaCl, 5.41mM KCl, 0.55 mM $CaCl_2$, 0.4 mM $MgSO_4$, 5.6 mM Na_2HPO_4, 0.035 M glucose, 0.04 mM phenol red, pH to 7.4
CIP	calf intestinal phosphatase
DAPI	4, 6-diamidino-2-phenylindole
DEPC	diethyl pyrocarbonate Used at 0.1% w/v to remove RNAase
DMSO	dimethyl sulphoxide
dNTP	deoxynucleoside triphosphate
FITC	fluorescein isothiocyanate
MMS	methylmethane sulphonate
MNE	MOPS/Sodium acetate/EDTA buffer 1x: 0.024 M MOPS, 5 mM NaOAc, 1 mM EDTA. PH adjusted to 7 with NaOH and stored in the dark at 4°C
NDS	solution for the manufacture of genomic plugs 1x: 0.5 M EDTA, 0.5M TRIS base, 0.5 M NaOH, 17 mM lauroyl sarcosine. pH adjusted to 8 or 9 with NaOH
PBS	phosphate buffered saline

PS	phosphate/sodium chloride buffer 1x: 0.06 M Na_2HPO_4, 46 mM NaCl.
PSG	phosphate/ sodium chloride/ glucose buffer 1x: 0.06 M Na2HPO4, 3.6 mM NaH_2PO_4, 46 mM NaCl, 55mM glucose, pH 8
RT	reverse transcriptase/transcription
SDS	sodium dodecyl sulphate
SOB	bacterial media (per Litre): 20 g bacto-tryptone, 5 g bacto-yeast extract, 0.5 g NaCl
SOC	SOB + 20 mM glucose
SSC	sodium chloride/ sodium citrate solution 1x: 0.15 M NaCl, 0.015 M $Na_3C_6H_5O_7$
TAE	TRIS/ acetate/ EDTA buffer 1x: 0.04 M TRIS base, 0.04 M glacial acetic acid, 1 mM EDTA
TBE	TRIS/ borate/ EDTA buffer 1x: 0.089 M TRIS base, 0.089 M ortho-boric acid, 2 mM EDTA
TB1/10E	TRIS/ borate/ 1/10 EDTA buffer 1x: 0.089 M TRIS base, 0.089 M glacial acetic acid, 0.2 mM EDTA
TE	10 mM Tris.HCl, 1 mM EDTA pH8.0

2.2 Trypanosome strains and their growth

The *Trypanasoma brucei* strain used in this study is 3174.2 (McCulloch *et al.*, 1997; Rudenko *et al.*, 1996), a transgenic derivative of the Lister 427 strain. *T. brucei* bloodstream form cells of strain Lister 427, expressing MITat1.2a (VSG 221), was derived from an unknown number of syringe passages through rodent hosts over many decades (Melville *et al.*, 2000). This strain switches the VSG being expressed at about 1 x 10^{-6} – 1 x 10^{-7} switches/cell/generation (see section 1.3), and is monomorphic, displaying only the long slender bloodstream form stage routinely. The transgenic 427 strain, 3174.2, contains hygromycin and G418 resistance genes in the ES containing *VSG 221* to allow VSG switching analysis (see section 4.8). *In vitro* growth used HMI-9 medium (Hirumi and Hirumi, 1989), whereas *in vivo* growth was carried out using adult female ICR mice (approx. 25 g) infected *via* interperitoneal injections.

2.3 Trypanosome isolation and stabilate preparation

Analysis of VSG switching and *in vivo* growth measurements required the recovery of trypanosomes from mouse blood after growth in the rodent host. For VSG switching (Section 2.4.1), the mice were exsanguinated by cardiac puncture, and the blood withdrawn into 5% sodium citrate anticoagulant in Carter's Balanced Salt Solution (CBSS) (0.15 ml CBSS/5% sodium citrate per 0.85 ml blood). To isolate clonal switched variants, 2 x 0.4 ml of exsanguinated mouse blood was subsequently centrifuged at 5000 rpm in a micro-centrifuge for 5 min. The centrifugation separates the blood into red blood cells as a bottom layer and plasma as a top layer, with *T. brucei* and white blood cells found in the interphase. The top and middle layers were removed with a needle (19 G) and syringe, added to 40 mls of HMI-9 and plated out over 2 x 96 well plates. For growth counts carried out during *in vivo* growth see Section 4.3.2.

Trypanosome stabilates were prepared by adding 100µl of 100% glycerol to 900µl of *T. brucei* cultures at a density of 1-2 x 10^6 cells.ml^{-1}, and 1 ml aliquots placed in cryotubes (Nunc). The stabilates were placed at -80°C overnight before being moved to liquid nitrogen for storage.

2.4 Transformation of the 3174.2 strain

T. brucei cultures, grown to a maximum density of 1-2 x 10^6 cells.ml^{-1}, were centrifuged at room temperature for 10 minutes at 1620 x g. The cells were then resuspended in 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 (ZMG) at a concentration of 1 x 10^8 cells.ml^{-1}. 5 x 10^7 trypanosomes were electroporated in 0.5 ml of ZMG at 1.5 kV and 25 μF capacitance using a Biorad Gene Pulser II. Approximately 5μg of DNA, that had been restriction digested, phenol/chloroform extracted and ethanol precipitated, was routinely used for all transformations. After electroporation cells were placed in 10 mls of HMI-9 for 24 hours before being placed on antibiotic selection. The recovered cells were centrifuged at 1620 x g for 10 minutes at room temperature and resuspended in HMI-9 (containing the appropriate drug) at a concentration of 10^7 cells per 18 mls. 1-2 x 10^7 cells (unless otherwise stated) were subsequently plated out in 24 well plates, with 10^7 cells plated out over 12 wells. Drugs used for selection were Hygromycin, G418, Blasticidin, Phleomycin and Puromycin at concentrations described in the text. Analysis of recombination efficiency used essentially the same procedure but different recovery periods and antibiotic selection concentrations (see Section 4.6 for details).

2.5 Analysis of VSG switching

The method used to analyse the frequency and mechanism of VSG switching is based upon that used by McCulloch et al (1997) and McCulloch and Barry (1999). Further details regarding this protocol can also be found in Chapter 4 (the analysis of switching frequency, Section 4.8.1, and the analysis of switching mechanism, Section 4.8.2).

2.5.1 Analysis of VSG switching frequency

To generate mice with acquired immunity against VSG221, mice were first injected with 2 x 10^5 3174.2 wild type cells that had previously been grown on hygromycin (5 μg.ml^{-1}) and G418 (2.5 μg.ml^{-1}) for a period of 5 days. These mice were then cured by injection of cymelarsan (Rhone Merieux; 5 mg.kg^{-1}) 3-4 days later. To generate

switched variants and calculate VSG switching frequency, the VSG221 immunised mice were then injected with 4-8 x 10^7 cells of the cell lines to be tested. These cell lines, that had previously been grown on hygromycin (5 $\mu g.ml^{-1}$) and G418 (2.5 $\mu g.ml^{-1}$) for a period of 5 days, were removed from antibiotic selection for 9 generations, thereby allowing VSG switch variants to arise. Approximately 24 hours after injection into the mice, the surviving *T. brucei* were recovered from the mouse blood by the buffy coat method and plated out as described in Section 2.2. The estimated switching frequency for each cell line was calculated from the number of wells displaying viable *T. brucei* populations as described in Section 4.8.1.

2.5.2 Analysis of VSG switching mechanism

The mechanisms of VSG switching that had been used in switched variants were determined using a combination of antibiotic selection and PCR (for further details see Section 4.8.2). 10 µl of cells from wells showing growth after recovery from VSG221 immunised mice were passaged into 1.5 mls of HMI-9 containing either hygromycin (5 $\mu g.ml^{-1}$), G418 (2.5 $\mu g.ml^{-1}$) or no drug. Up to 10 days later, the cells were scored for drug resistance and genomic DNA prepared (as described in section 2.6) from the cells grown on no drug. PCR analysis was then carried out on the genomic DNA to determine the presence or absence of the hygromycin and G418 resistance genes, and the VSG221 gene as described in section 4.8.2.

2.6 Isolation of genomic DNA

Trypanosomes were harvested from culture by centrifugation at 1620 x g for 10 min at room temperature and resuspended in digestion buffer (50mM Tris.Cl (pH 8.0), 1mM EDTA, 100mM NaCl). 50 µl 10% SDS and 2.5 µl of proteinase K (at 20 $mg.ml^{-1}$) were then added, and the preparations were incubated at 37°C overnight to lyse the trypanosomes and digest proteins. The DNA was recovered from the lysis reaction by phenol/chloroform extraction and ethanol precipitation (see Section 2.6.1).

2.6.1 Phenol/chloroform extraction and ethanol precipitation

To extract genomic DNA from the *T. brucei* lysis reactions, or from other reaction mixtures, such as restriction digestions, an equal volume of phenol/chloroform (at a 1:1 mixture) (Sigma) was added and the solution mixed thoroughly by gentle inversion (for genomic DNA) or vortexing (for other DNA forms). The two phases were separated by centrifugation at maximum speed in a microcentrifuge for 3 min at room temperature, after which the aqueous layer was removed and transferred to a fresh eppendorf tube. 2 volumes of 100% ethanol and $^{1}/_{10}$ the volume of 3 M sodium acetate (pH 5.2) were then added and the tube contents were mixed thoroughly. The tube was then incubated at -20°C for a minimum of 20 min, after which the DNA was harvested by spooling (for genomic DNA) or pelleted by centrifugation at maximum speed in a microcentrifuge for 30 min at 4°C. The pellet was then washed in 1 ml 70% ethanol, air-dried, and resuspended in an appropriate volume of buffer (usually TE) or water.

2.7 Isolation of total RNA

Trypanosomes were harvested from culture medium by centrifugation at 1620 x g for 10 min at room temperature. Total RNA was then isolated using the Qiagen RNeasy mini kit, following the manufacturer's protocol and using a 25 G needle and a syringe to lyse the cells.

2.8 Polymerase chain reaction (PCR)

PCRs were set up in either 25 or 50 µl volumes. Routinely, diagnostic PCRs were set up in a 25 µl volume, whereas, PCRs to amplify DNA fragments for cloning were set up as 50 µl reactions. 50 µl reactions were prepared as an exact double of the 25 µl reactions, and therefore only the setting up of the 25 µl reactions is described. For each reaction 0.25 µl (1.25 units per reaction) of either Taq (ABgene – 5 $U.\mu l^{-1}$), or Herculase (Stratagene – 5 $U.\mu l^{-1}$), DNA polymerase was commonly used to amplify DNA. For each polymerase, 2.5 µl of supplied 10 x buffers and the manufacturer's recommended $MgCl_2$ was added, along with 1 µl of 10 mM dNTPs and 1 µl of each primer at 5 mM. The reaction conditions used were 95°C for 5 mins, followed by 30 cycles of 95°C for 1

min, 50-60°C for 1 min and 72°C for 1 min per kb, and finally 72°C for 10 mins. PCR machines used were for the amplification of DNA were a Robocycler® (Stratagene) and a PCRsprint (Hybaid). PCRs were routinely purified using the Stratagene PCR purification kit, following the manufacturer's protocol. Appendix 1 contains a list of the oligonucleotides used for PCRs.

2.8.1 Reverse transcription polymerase chain reaction (RT-PCR)

Before cDNA preparation, *T. brucei* RNA was treated with DNAase to remove any contaminating genomic DNA. 2 µg of RNA was incubated, for 15 min at room temperature, with 1 unit of DNAase (Invitrogen) and 1 µl of 10 x DNAase buffer, in a total volume of 10 µl made up with water. The reaction was terminated by the addition of 1 µl of 0.25 mM EDTA and heating to 65°C for 20 min.

cDNA was then prepared from the DNAase treated RNA using the SuperscriptTM First-Strand Synthesis System for RT-PCR kit (Invitrogen). 50 ng of random hexamers were added to 8 µl of the DNase treated RNA, with 1 µl of 10 mM dNTPs, and heated to 65°C for 5 min and then incubated on ice for 1 min. 4 µl 25mM $MgCl_2$, 2 µl 0.1 M DTT, 2 µl of 10 x RT buffer and 1 µl of RNaseOUT recombinant ribonuclease inhibitor were added to the RNA solution and incubated for 2 min at 25°C. 1 µl of Superscript II reverse transcriptase (200 $U.ml^{-1}$) was then added, and the reaction was incubated at 25°C for 10 mins before being incubated at 42°C for 50 mins. For each cDNA preparation an identical reaction was also set up with the same DNAase treated RNA, but with water added instead of reverse transcriptase. This acted as a negative control for the presence of DNA contamination in subsequent PCRs. Following cDNA generation, reverse transcriptase was heat-inactivated at 70°C for 15 min and the tube was subsequently chilled on ice. To remove any remaining single-stranded RNA, 1 µl of RNAase H (3.8 $U.ml^{-1}$) was added and the reaction was incubated at 37°C for 20 min. This cDNA was then suitable as a template for PCR, with 1 µl commonly used in 25 µl PCRs.

2.8.2 3′ rapid amplification of cDNA ends (RACE)

To obtain the 3′ sequences of partially sequenced genes, 3′ RACE was carried out on cDNA generated from total RNA using the 3′ RACE System for Rapid Amplification of cDNA Ends kit (Life Technologies). The manufacturers protocol was followed with the exception of the steps relating to the generation of cDNA that were missed out as cDNA, generated using random oligonucleotides, and not RNA, was used as a starting template (further details regarding this protocol can be found in section 3.3).

2.9 DNA digestion, electrophoresis and Southern blotting

2.9.1 Restriction enzyme digestion of DNA

Restriction enzyme digestion of DNA was performed at the specified temperature using commercial restriction enzymes. Routinely 1 to 2 μg of DNA was digested in a reaction volume of 20 μl with 1 to 2 U of restriction enzyme (NEB) and 2 μl of the recommended buffer (10x Buffer, NEB). If a larger quantity of digested DNA was required, the reactions were scaled up, to a maximum volume of 100 μl, and were subsequently phenol:chloroform extracted (Section 2.6.1). Plasmid DNA was routinely digested for 2 hours, whereas, genomic DNA was digested for a period ranging from 6 hours to overnight.

2.9.2 Gel electrophoresis

Standard DNA separations were performed on 1.0% agarose gels (Seakem LE, BMA) run at 100V in 1 x TAE buffer using a commercial 1 kb ladder as a size marker (Invitrogen). Digests of genomic DNA for Southern blots were electrophoresed on 0.8% agarose gels (Seakem LE, BMA) run at 30 V for 16-24 hours in 1 x TBE buffer. In both cases, gels routinely contained 0.2 $\mu g.ml^{-1}$ ethidium bromide (EtBr) to facilitate visualisation of the DNA under UV light.

2.9.3 Southern blotting

Agarose gels to be Southern blotted were initially photographed on a UV transilluminator with a ruler parallel to the gel so that the sizes of bands detected by the hybridisation of radioactively labelled DNA (see section 2.10) could be measured. To nick the DNA, the gel was placed in 0.25 M HCl for 15 min and then rinsed with distilled water. Following this, the DNA was denatured by placing the gel in denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 30 min. After rinsing with distilled water, the gel was placed in neutralising solution (1 M Tris-HCl pH 8.0, 1.5 M NaCl) for a further 30 min. The DNA was transferred to a nylon membrane (Hybond-XL) by capillary blotting using 20 x SSC as the transfer buffer (Sambrook et al., 1989). Blotting was routinely performed for 24 h and then the DNA was crosslinked to the membrane using the auto crosslink setting on a UV stratalinker® (Stratagene).

2.10 Pulse Field Gel Electrophoresis (PFGE)

Chromosome sized DNA was separated by PFGE using the CHEF-DR III system (BIO-RAD).

2.10.1 The preparation of genomic plugs

Each genomic agarose plug contained 5 x 10^7 trypanosomes. The trypanosomes, growing in HMI-9, were centrifuged at 1620 x g for 10 min at room temperature and washed twice in the same volume of PSG. For a single agarose plug, 5 x 10^7 *T. brucei* cells were then centrifuged at 1620 x g for 10 min at room temperature, the pellet resuspended in 50 µl of PSG and warmed to 37°C for 1 min. This was then mixed with 50 µl of 2% low-melting point agarose (InCert agarose, FMC Bioproducts) solution (which had been prepared in 1 x $TB^1/_{10}E$ and held at 37°C), and dispensed into a plug mould (BIO-RAD) and left to set at 4°C.

When solid, plugs were dispensed into tubes to which a minimal volume of NDS pH9 containing 1 $mg.ml^{-1}$ proteinase K was added, and the tubes incubated at 50°C for 24 hours. The buffer was then changed to NDS pH8 containing 1 $mg.ml^{-1}$ proteinase K and

incubated at 50°C for another 24 hours. After this, the plugs were transferred to NDS pH8 for storage at 4°C.

2.10.2 Pulse Field Gel Electrophoresis

3 litres of 1 x $TB^1/_{10}E$ was placed in the PFGE tank and allowed to circulate at the running temperature of 12°C. After 30 mins this buffer was used to equilibrate the genomic plugs for 1 hour at room temperature and then stored at 4°C overnight. The following day, 110 mls of 1 x TB1/10E was removed from the PFG tank and used to make a 1% agarose gel (Seakem Gold agarose, FMC Bioproducts). 100 mls of this was poured into a gel tray and the remainder kept at 37°C. When solid, the comb was removed, the agarose genomic pugs placed into the wells and the wells sealed using the remaining agarose. The gel was then electrophoresed at 4.6 $V.cm^{-1}$ for 42 hours with the electrode switching time linearly ramped from 8 – 15 seconds. The gel was then stained by soaking in 200 mls of the running buffer containing 4µl of EtBr (10 $mg.ml^{-1}$), and then rinsed with distilled water before visualisation under UV light. PFGE gels were then Southern blotted as described in Section 2.8.3.

2.11 RNA electrophoresis and Northern blotting

2.11.1 RNA electrophoresis

RNA separations were performed on 1.0% agarose gels (Seakem LE, BMA) containing 2.2M formaldehyde, and run at 30 V for 16-24 hours in 1 x MNE buffer using a commercial 0.2-10 kb ladder as a size marker (Sigma). The RNA samples were added to 20µl of RNA loading buffer (60 µl formaldehyde, 20 µl formamide, 24 µl 5 x MNE buffer and 16 µl of H_2O) and 1µl of EtBr (0.2 $\mu g.ml^{-1}$) and incubated at 65°C for 5 mins before separation.

2.11.2 Northern blotting

Agarose gels to be Northern blotted were initially photographed on a UV transilluminator with a ruler parallel to the gel so that the sizes of bands detected by the hybridisation of radioactively labelled DNA (see section 2.10) could be measured. Gels were then soaked in sodium phosphate (pH 6.5) for 15 mins to remove the formaldehyde. The RNA was transferred to a nylon membrane (Hybond-XL) by capillary blotting using sodium phosphate (pH 6.5) as the transfer buffer. Blotting was routinely performed for 24 h and then the RNA was crosslinked to the membrane using the auto crosslink setting on a UV stratalinker® (Stratagene).

2.12 Radiolabelling and hybridisation of DNA probes

2.12.1 Probe manufacture by random hexamer radiolabelling of DNA

The DNA fragments used for probes in this study were PCR products that were gel extracted using the Qiagen gel extraction kit following the manufacturers protocol. Radiolabelling of these DNA fragments was performed using a commercial kit (Prime-It II kit, Stratagene). 25 ng of the DNA template was mixed with 10 µl of random-sequence oligonucleotides and sterile, distilled water, in a total reaction volume of 36 µl. The mixture was then heated to 95°C for 5 min to denature the DNA and cooled to allow the random oligomers to anneal. 10 µl 5 x primer buffer, 3 µl of ^{32}P labelled dCTP (30 µCi) and 1 µl Klenow (5U. μl^{-1}) was then added, mixed, and incubated at 37°C for 10 min. The resultant probes were then purified from the unincorporated nucleotides by passage through Microspin columns (Amersham) by centrifuging at 3000 rpm in a microcentrifuge for 2 mins at room temp. After purification, the probes were denatured at 95°C for 5 min before hybridisation.

2.12.2 Hybridisation of radiolabelled DNA probes

Nylon filters of blotted DNA or RNA (Sections 2.8.3 and 2.10.2) were placed in glass hybridisation tubes (Hybaid) and approximately 50 mls of pre-warmed Church-Gilbert solution (0.342 M Na_2HPO_4, 0.158 M $NaH_2PO_4.2H_2O$, 0.257 M SDS and 1 mM EDTA

per litre) added and the filters pre-hybridised for a minimum of 1 h at 65°C in a rotating hybridisation oven. The denatured, radiolabelled probe (Section 2.11.1) was then added to the Church-Gilbert solution in the tube, and the hybridisation was left for 16-24 h at 65°C. After hybridisation, the filters were washed at 65°C, in a rotating oven, with 2 x SSC, 0.1% SDS for 30 mins and then 0.2 x SSC, 0.1% SDS for another 30 mins. The filters were finally rinsed in 0.1 x SSC at room temperature and heat-sealed in plastic. The membrane was visualised using a phosphoimage screen (Fuji) at room temperature for 2 to 48 h and developed using a Typhoon 8600 (Amersham), again depending on the strength of the signal.

2.12.3 Stripping of hybridised nylon-membranes

Nylon membranes of blotted DNA or RNA were stripped of hybridised probe fragments by pouring boiling 0.1% SDS onto them whilst present in a heat resistant container and then allowing the SDS to cool to room temperature. The procedure was repeated again, and the membrane rinsed in 0.1 x SSC. Following stripping, the filters were checked for successful stripping using a phosphoimage screen (Fuji) at room temperature for 24 h. Following this, the membranes were then ready for re-hybridisation.

2.13 Cloning of DNA fragments

2.13.1 Cloning using T4 DNA ligase

DNA fragments for cloning were PCR amplified and PCR purified (Section 2.7) and restriction digested (Section 2.8.1). On occasion, DNA fragments for cloning required an overhang to be filled in prior to ligation. This was achieved for 5′ overhangs by using the Klenow fragment of *E. coli* polymerase I (NEB), after heat-inactivation of the restriction enzyme(s) used to cleave the DNA. 1 µl of klenow, 1 µl of 10 mM dNTPs, 0.5 µl of the same restriction enzyme buffer (NEB), and 2.5 µl of H_2O were added to the heat-inactivated digestion reaction, and incubated at 37°C for 30 min. Filling in of 3′ overhangs was not carried out during this project.

Self-ligation of the restriction digested vectors was prevented by treatment with Calf-intestinal phosphatase (CIP; Roche Diagnostics), which removed the 5′ phosphate

groups. 0.5 µl of CIP (10 $U.\mu l^{-1}$) was added to the digestion reaction, which was incubated at 37°C for 10 min. After CIP treatment, the vector was purified either by phenol:chloroform extraction or gel extraction (using the Qiagen gel extraction kit) following electrophoresis.

Ligation of DNA fragments into a plasmid vector (e.g. pBluescript (Stratagene) derivatives thereof) were carried out in a 10 µl volume with 1 µl. of T4 DNA ligase (400 $U.ml^{-1}$; NEB) and 1 µl of ligase buffer (NEB) which was incubated at room temperature for 6 hours. 6µl of this ligation reaction was subsequently used to transform 80µl of E. coli Xl-1 blue MRF′ cells (see section 2.11.3).

2.13.2 Cloning into TOPO vector

PCR products generated by use of Taq DNA polymerase (AB gene) were suitable for cloning into this vector, as this enzyme produces a 3′ single adenosine overhang. 0.5 – 4 µl of PCR product was incubated with 1 µl of salt solution and 1 µl of TOPO TA vector (Invitrogen), in a total volume of 6 µl made up with water, for 5 min at room temperature. 2 µl of this reaction was subsequently used to transform 25µl of TOP 10F′ (Invitrogen) cells (see section 2.11.3).

2.13.3 Transformation of *E. coli* and plasmid retrieval

The transformation of both MRF′ and TOP 10F′ *E. coli* cells was carried using heat shock. Ligations and cells were mixed gently and left on ice for 30 min. Following this, the cells were heat shocked at 42°C for 45 s, and then immediately transferred to ice for 2 mins. Transformations of ampicillin resistance containing plasmids were then spread over L-agar plates containing ampicillin at a final concentration of 50 $\mu g.ml^{-1}$ (Sigma). However, transformations of chloramphenicol resistance containing plasmids required expression of the drug resistance gene before plating on selective media. To do this, 1 ml of SOC was added to the transformed cells, which were then incubated at 37°C for 1 h. The cells were then centrifuged for 1 min at 10,000 rpm in a micro-centrifuge, the supernatant was poured off, and the pellet resuspended in 100 µl of SOC. This suspension of cells was then spread over L-agar plates containing chloramphenicol at a final concentration of 25 $\mu g.ml^{-1}$ (Sigma), and incubated overnight at 37°C. Single

colonies from these plates were then picked and used to inoculate L-broth (also containing the appropriate antibiotic at the same concentration as contained in the agar plates), and grown up overnight at 37° C. Plasmids were then prepared from 1.5-3 ml of the overnight culture using the Qiagen miniprep kit, or from 200 - 400 ml of culture using the Qiagen Maxiprep kit, following the manufacturers instructions.

2.14 *T. brucei* staining and hybridisation

T. brucei cells were visualised using an Axioskop 2 microscope (Zeiss) and Openlab software (Improvision).

2.14.1 DAPI staining

1 ml of *T. brucei* culture, at a density of 1-2 x 10^7 cells.ml^{-1}, was centrifuged, washed twice with PBS before being resuspended in 1 ml of PBS. 10 µl samples were then spotted onto multi-spot microscope slides (C.A.Hendley Ltd.) and allowed to air-dry. The trypanosomes were then fixed by soaking in methanol for 5 minutes at room temperature and again allowed to air dry before vectashield with DAPI (Vector Laboratories Inc.) was added, a coverslip positioned and the slide sealed with clear nail varnish.

2.14.2 *in situ* hybridisation

The *in situ* hybridisation protocol is described in Section 5.2.2. The RAD51 antibody used in this assay was a rabbit polyclonal antiserum (Diagnostics Scotland) generated in response to His tag purified, *E. coli* expressed recombinant *T. brucei* RAD51 (supplied by K. Norrby). FITC conjugated secondary antibody used was a goat-derived anti-rabbit IgG (whole molecule; Sigma).

CHAPTER 3

RAD51-LIKE GENES IN *T. BRUCEI*

3.1 Introduction

Humans, yeast and bacteria are well known to have an array of proteins that act to repair DNA and maintain genomic integrity. This array of proteins revolves around a core strand exchange enzyme, Rad51 (RecA in bacteria), and a number of cofactors that aid Rad51 in its role as a repair enzyme. So far only RAD51 has been shown to have a role in recombination during antigenic variation in *T. brucei*. However, in a *RAD51* mutant repair, recombination and antigenic variation still occur, albeit at a reduced level. As a result it was concluded that back-up pathways must exist for these processes. To examine this question, this project has attempted to identify potential *RAD51*-like genes in *T. brucei*, since these are well-characterised examples of factors that act in homologous recombination. To do this, the *T. brucei* genome was BLAST searched with the amino acid sequences of *S. cerevisiae* and *T. brucei* RAD51, and with *E. coli* RecA (see Appendix 2 for accession numbers). As a result of this, five further *RAD51*-related genes were identified, one of which was incomplete at the outset of this work. This chapter describes the experiments carried out to obtain the missing 3′ sequence of one of the *RAD51*-like genes, and sequence analysis to examine the potential for each of the *RAD51*-related genes to encode a strand exchange enzyme. In addition, this chapter describes initial functional characterisation of three of the *RAD51*-like genes.

3.2 Identification of putative *RAD51*-like genes in *T. brucei*

The *T. brucei RAD51*-like genes were identified through BLAST searches of the incomplete (at the outset of this work) *T. brucei* database, carried out throughout the course of this project. TBLASTN searches of sequence reads and assembled contigs were carried out using *S. cerevisiae* and *T. brucei* RAD51, and *E. coli* RecA, as query sequences, searching via omniblast at Gene DB (Sanger). At the start of the project, Oct/Nov 2001, RAD51 searches revealed the complete *RAD51* sequence, a complete lowly related sequence (subsequently named *RAD51-3*) and a partial, highly related sequence (section 3.3; subsequently named *DMC1*). Searches with *E. coli* RecA revealed a complete, highly divergent open reading frame not initially recognised by the RAD51 searches (subsequently named *RAD51-5*). The same searches revealed multiple further small sections of homology that were monitored through out to see if they became part of larger contigs containing other genes. Later in the project, approximately

Dec 2002, this revealed two further lowly-related genes (subsequently named *RAD51-4* and *RAD51-6*).

3.3 3′ rapid amplification of cDNA ends to obtain *DMC1* 3′ sequence

Before the analysis of one of the *RAD51*-like genes, named *DMC1* (see below), could be conducted, the complete open reading frame needed to be established, as the 3′ end was absent in the available *T. brucei* genome sequence. At the beginning of this project, the incomplete *T. brucei* genome database contained only 780 bp of *DMC1*. To obtain the remainder of the sequence, 3′ rapid amplification of cDNA ends (3′ RACE, Life Technologies) was carried out. This technique utilises the poly-A tail, added to the 3′ end of every mRNA, to anneal an adapter primer in order to generate cDNA from mRNA sequences in total RNA extracts. Reverse Transcriptase-PCR (RT-PCR) is then carried out to amplify the gene of interest using a gene-specific primer and an 'abridged universal amplification primer', specific for the adapter primer (Fig 3.1).

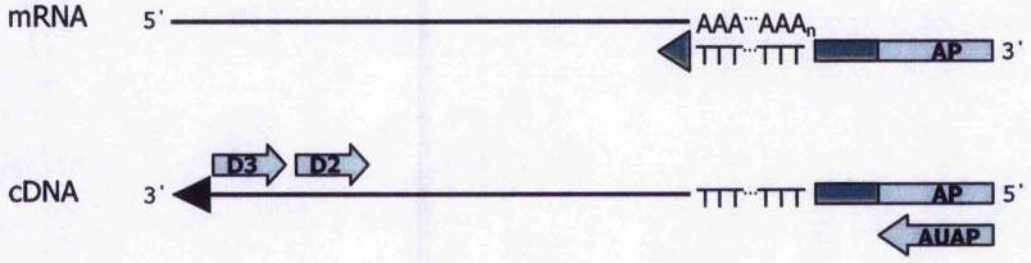


Figure 3.1: The use of 3′ RACE to amplify the 3′ end of genes. The adapter primer (AP) is used to generate cDNA by reverse transcription from total RNA, after which the RNA is degraded. The 3′ end of a gene of interest can then be amplified by RT-PCR using a gene specific primer, in this case *DMC1*-D3 (D3) or *DMC1*-D2 (D2), and the abridged universal amplification primer (AUAP), specific for the adapter primer.

For the *DMC1* gene, RT-PCR was carried out on cDNA generated, as described above, from Lister 427 bloodstream form-derived total RNA (A. Lewis, gift) using Taq DNA polymerase and primers *DMC1*-D3, *DMC1*-D2 (Appendix 1) and the abridged universal amplification primer (AUAP). The primers *DMC1*-D3 and *DMC1*-D2 were designed using sequence 231 bp and 146 bp, respectively, from the 3′ end of the available *DMC1* sequence. RT-PCR reactions were carried out using both combinations of primers, as

well as individual primer controls, from both RT-treated and RT-untreated cDNA samples (Fig 3.2).

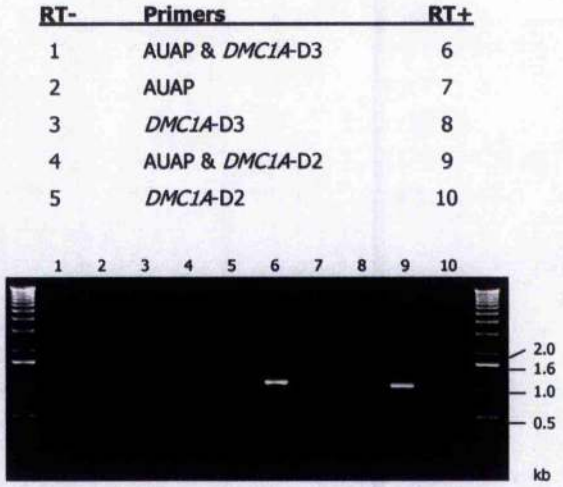


Figure 3.2: 3′ RACE reactions carried out to obtain *DMC1* 3′ sequence. RT-PCR reactions were carried out as listed, of which 5 µl samples were separated by agarose gel electrophoresis. Primer names are shown in Fig 3.1 and discussed in the text. RT+: reverse transcriptase treated cDNA. RT-: reverse transcriptase untreated cDNA.

A strong product was generated in each of the two-primer reverse transcriptase plus reactions, with AUAP and *DMC1*-D3 (Fig 3.2, lane 6) resulting in a PCR product approximately 1150 bp in size, and AUAP and *DMC1*-D2 (Fig 3.2, lane 9) resulting in a PCR product approximately 1050 bp in size. The size difference between the two products, approximately 100 bp, suggested that the products are *DMC1*-specific, as the two gene-specific primers hybridise 85 bp apart. Furthermore, no products were generated in the reverse transcriptase untreated controls, and all single primer reactions gave diffuse products, even though both the *DMC1*-D3 and D2 primers (Fig 3.2, lanes 8 and 10) produced fragments of similar size to their cognate two primer reactions (Fig 3.2, lanes 6 and 9). The two specific products were Topo cloned (TOPO TA cloning kit, Life Technologies; section 2.12.2) and sequenced (MBSU sequencing service, Glasgow). The sequences obtained from each product and the existing *DMC1* sequence were then aligned using Contig Express (Vector NTI suite 9, Invitrogen life science software) resulting in the assembly of an uninterrupted open reading frame 1047 bp in length. Almost immediately after this derivation of the putative *DMC1* open reading frame, the *T. brucei* genome database was updated and included the complete *DMC1* sequence. Comparison of the *DMC1* sequence obtained by 3′ RACE with that contained within the database was carried out using AlignX (Vector NTI suite 9, Invitrogen life

science software), which showed the sequences to be near identical (data not shown). For the analysis of the RAD51-like genes during this project gene sequences from the T. brucei genome database were used.

3.4 Sequence homologies and phylogenetics

Having identified 5 further *RAD51*-related genes by searching the *T. brucei* genome database, BLASTp searches were carried out against the NCBI database using the predicted amino acid sequences of each of the genes (Table 3.1). This provided a crude estimate of whether or not these proteins are identifiable as Rad51- or RecA-like proteins in other organisms, rather than other proteins. In all cases the most recognised proteins were indeed Rad51- or RecA-related proteins, although for two (RAD51-5 and RAD51-6) the level of confidence in these hits was very low. The genes were given generic names based on their homology to Rad51 (RecA) sequences, although this was not intended to infer a prediction of their function.

Gene	Homologue	BLASTp Hit
RAD51-3	*Homo sapiens* Rad51C	2.00E-27
	Chlamydomonas reinhardtii Rad51C	3.00E-25
	Mus musculus Rad51C	4.00E-25
RAD51-4	*Schizosaccharomyces pombe* DMC1	5.00E-12
	Coprinopsis cinerea Lim15/DMC1	1.00E-10
	Entamoeba histolytica DMC1	1.00E-10
RAD51-5	*Acidocella facilis* RecA	0.002
	Vibrio cincinnatiensis RecA	0.007
	Aeromonas salmonicida RecA	0.007
RAD51-6	*Arabidopsis thaliana* XRCC3β	2.00E-07
	Methanococcoides burtonii RecA/RadA	0.005
	Zea mays Rad51A	0.035
DMC1	*Leishmaniama major* DMC1	1.00E-145
	Oryza sativa DMC1B	1.00E-113
	Oryza sativa Lim15B	1.00E-113

Table 3.1: BLASTp search results for the *T. brucei RAD51* homologues. Three of the top hits obtained when carrying out BLASTp searches using the amino acid sequences of the *T. brucei RAD51*-related genes are shown.

The predicted amino acid sequences for each of the 5 *RAD51*-related genes were next analysed to determine their potential to encode strand exchange enzymes. To do this, the sequences were compared to the conserved domains of strand exchange enzymes identified in the study by Brendel *et al* (1997). In this work, 15 Rad51/Dmc1/RadA proteins and 13 RecA proteins from a diverse range of organisms were compared to define conserved domains within this group of proteins. The authors identified 6 domains common to each of the Rad51/Dmc1/RadA sequences analysed and within those domains, eight sub-domains conserved also within the RecA sequences (Fig 3.3).

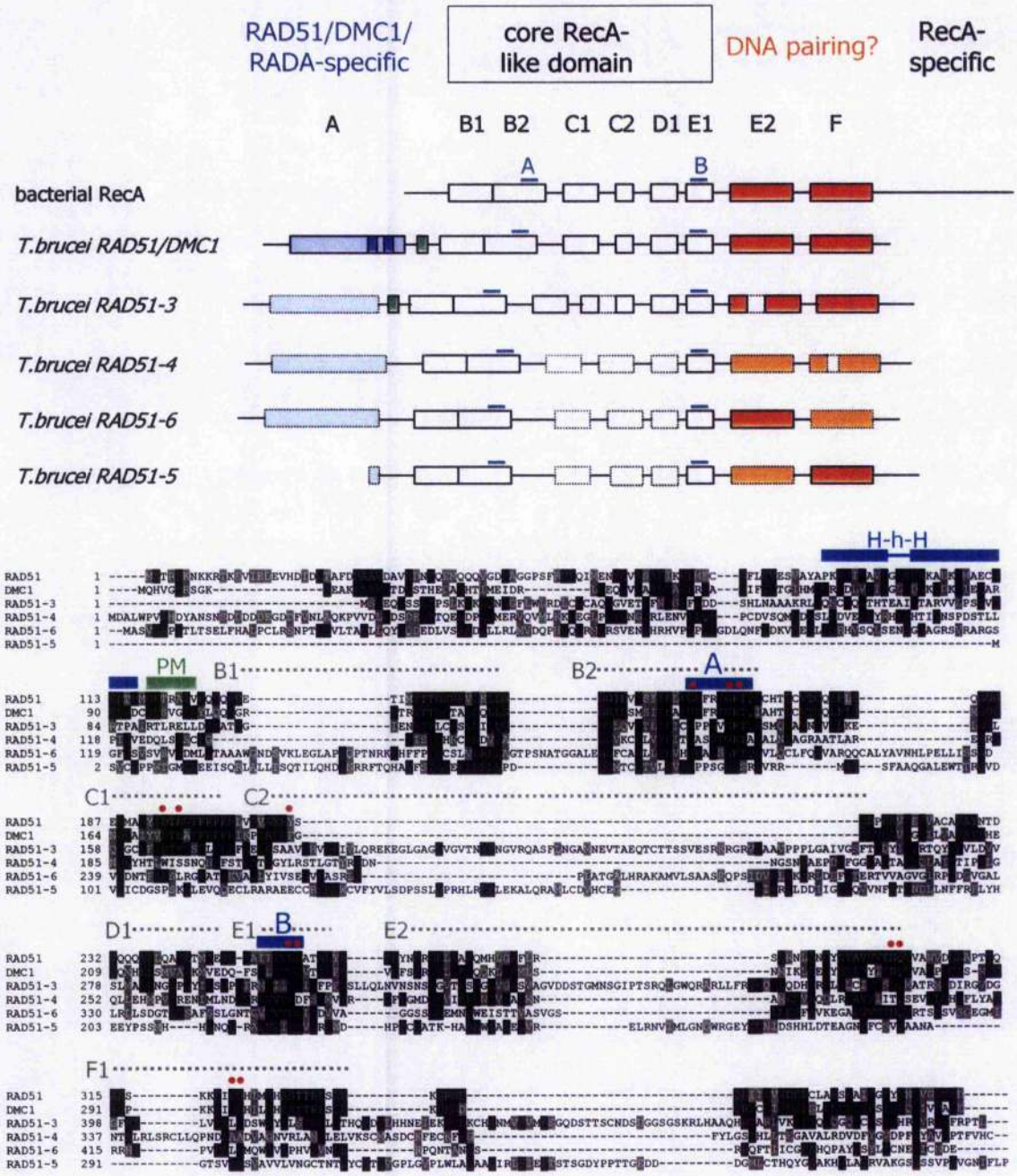


Figure 3.3: Alignment of the amino acid sequences of *T. brucei* RAD51 and the RAD51 homologues. Top: The alignment of the polypeptide sequences of the *T. brucei* RAD51-related genes based on the domains identified by Brendel *et al* (1997) with general domain functions indicated. A: possibly interacts with other factors and/or nucleic acids. B1 & B2: encompasses the hydrophobic core of the RecA structure and contains the Walker A box that binds ATP. C1: comprises a highly conserved monomer-monomer surface. C2: conserved between the Rad51/Dmc1/RadA and RecA sequences. D1: comprises a helix in the RecA structure and forms part of a monomer-monomer surface. E1: contains the Walker B box that hydrolyses ATP. E2: forms part of 2 loops, a helix and a β-strand. F: involved in monomer-monomer interactions and DNA binding. RecA-specific: possibly binds LexA/UmuD. Lighter colours and dashed lines represent lower conservation. **Bottom:** *T. brucei* RAD51 polypeptide sequence (top) is compared to all of the *RAD51*-like genes found in *T. brucei* lined up using the conserved Walker A (A) and B (B) boxes. See appendix 2 for accession numbers. The conserved domains described by Brendel *et al* (1997) are shown (grey writing and dashed line) as are the putative BRC consensus motif (PM; Pellegrini *et al*., 2002), the helix-hairpin-helix motif (H-h-H; Grishin, 2000) and the conserved residues (•; Story *et al* (1993). The alignment was carried out using ClustalW (Chenna *et al*., 2003) and coloured using Boxshade 3.21 (http://www.ch.embnet.org/software/ BOX_form.html).

In addition, Rad51/Dmc1/RadA sequences were found to contain an N-terminal extension (domain A) absent from the bacterial RecA sequences, and the RecA sequences were found to contain a unique C-terminal extension. The 5 recently identified *T. brucei* RAD51-like polypeptides and the canonical *T. brucei* RAD51, were aligned with each other by ClustalW, via Vector NTI (Vector NTI suite 9, Invitrogen life science software) and manual adjustments made to the alignment based on the conserved residues identified by Brendel *et al* (1997). From the sequence comparisons most of the conserved domains described by Brendel *et al* could be identified in each of the proteins. All of the conserved domains could be unambiguously identified in *T. brucei* RAD51 and DMC1. *T. brucei* RAD51-3 appears somewhat more diverged, with putative insertions in sub domains C2 and E2 and sequence divergence in the N-terminal A domain. *T. brucei* RAD51-4, RAD51-5 and RAD51-6 are yet further diverged. Sub domains A, C1, C2, D1, E2 and F were difficult to identify in RAD51-4. Similarly, A, C1, C2, D1 and F were diverged in RAD51-6. Finally, RAD51-5 was diverged in C1, C2, D1 and E2 and appeared to be severely truncated in the Rad51/Dmc1/RadA specific N-terminal A domain, assuming the predicted start codon is correct. Despite the above sequence divergences, putative Walker A and B boxes (Walker *et al*., 1982), responsible for ATP binding and hydrolysis, could be identified in all the *T. brucei* RAD51-like sequences, with one possible exception. In RAD51-4, a conserved serine residue has been changed to aspartic acid within the putative Walker B box (D277). Assuming this is a genuine coding change and not a sequence error, it suggests that all the *T. brucei* RAD51-like proteins, except RAD51-4, are capable of ATP binding and hydrolysis.

Story *et al* (1993) identified 12 amino acid residues central to the catalytic activity and structural integrity of Rad51, Dmc1 and RecA in many organisms. The sequence comparisons suggest that although these are universally conserved in *T. brucei* RAD51 and DMC1, they are much less conserved in the four other proteins. This may indicate that these are not capable of strand exchange themselves, but contribute to RAD51 or DMC1 catalysis. However, no 3-dimensional structure for any RAD51-related protein, other than DMC1 (Masson *et al*., 1999; Passy *et al*., 1999), has been determined, so this is a preliminary conclusion. Further sequence elements with Rad51 proteins have been determined by recent structural studies (Pellegrini *et al*., 2002; Shao and Grishin, 2000). Pellegrini *et al* (2002) in an attempt to define the interactions between Rad51 and Brca2 analysed Rad51 sequences looking for motifs similar to the BRC repeat motifs. They found a highly conserved BRC repeat motif conserved within Rad51 polypeptides from

a range of organisms. This motif could only be found in *T. brucei* RAD51 and was absent from all the other RAD51-related proteins. *T. brucei* DMC1 sequence contained a highly conserved glycine residue in this motif, but the conserved alanine has changed to a glycine (G99) and it was deemed that this would result in a non-functional motif. Shao and Grishin (2000) describe helix-hairpin-helix motifs, which are involved in non-sequence-specific DNA binding, present within a wide range of proteins. In this work they show two conserved helix-hairpin-helix motifs, linked by an α-helix, conserved in the N-terminal region of *H. sapiens* Rad51. Analysis of the polypeptide sequences of the *T. brucei RAD51*-related genes was carried out to determine the presence or absence of helix-hairpin-helix motifs. From these results, RAD51 and DMC1 appeared to contain one helix-hairpin-helix motif, whereas they appeared to be absent in all of the other RAD51-related proteins. All conserved helix-hairpin-helix residues found in *H. sapiens* Rad51 were found to be conserved in *T. brucei* RAD51; however, in DMC1 the central conserved residue has been changed to leucine (L76), although this reside was conserved in the helix-hairpin-helix motifs in other proteins.

The levels of sequence homology of the *T. brucei* RAD51-like polypeptides, compared with *T. brucei* RAD51 or DMC1, in the alignments appear to be somewhat low. This result was confirmed by determining the sequence identities of the *T. brucei* RAD51-like polypeptides to *T. brucei* RAD51 in pairwise comparisons (Table 3.2). To examine this question further, phylogenetic analysis was carried out using the polypeptide sequences of a number of RecA and Rad51-like proteins from a wide range of organisms (Fig 3.4). To do this, only those Rad51 or Rad51-related proteins that have been functionally examined were considered. These were *T. brucei RAD51* (McCulloch and Barry, 1999), *Leishmania major RAD51* (McKean *et al.*, 2001), *Plasmodium falciparum RAD51* (Bhattacharyya and Kumar, 2003), *Saccharomyces cerevisiae RAD51/55/57* and *DMC1* (Paques and Haber, 1999), *Schizosaccharomyces pombe Rhp51/55/57* (Grishchuk and Kohli, 2003) and *Dmc1* (Fukushima *et al.*, 2000), *Homo sapiens RAD51/L1/C/L3, XRCC2/3* (Masson *et al.*, 2001b) and *DMC1* (Masson *et al.*, 1999), *Drosophila melanogaster Rad51/51C/51D* and *SpnB/D* (Staeva-Vieira *et al.*, 2003), *Arabidopsis thaliana Rad51/51C, XRCC3* and *Dmc1* (Osakabe *et al.*, 2002), *Oryza sativa DMC1A/B* (Kathiresan *et al.*, 2002), *Ustilago maydis Rad51 and Rec2* (Ferguson *et al.*, 1997), *Escherichia coli RecA* (Brendel *et al.*, 1997), *Bacillus subtilis RecA* (Brendel *et al.*, 1997), *Streptomyces lividans RecA* (Brendel *et al.*, 1997), *Campylobacter jejuni RecA* (Guerry *et al.*, 1994), *Neisseria gonorrhoeae RecA* (Koomey and Falkow, 1987) and *Staphylococcus aureus RecA* (Bayles *et al.*, 1994).

Included in the phylogenetic comparisons were the *T. brucei RAD51*-like genes, and *Dmc1* from *L. major* and *P. falciparum*, none of which have been functionally characterised. In this analysis, the Rad51, Dmc1 and RecA proteins formed discrete groupings, whereas the Rad51-like proteins showed less similarity and did not group together. This suggests that the level of conservation, beyond the Rad51 and Dmc1 polypeptides, in eukaryotes is overall low, and it is not possible to infer functional relationships between the RAD51-like proteins. Whether or not this means that these genes arose independently in each organism, or have loosely defined functions that allow broad sequence diversification, is not clear (Bennett and Holloman, 2001; DiRuggiero *et al.*, 1999; Thacker, 1999; Yang *et al.*, 2001).

To illustrate this further, all *RAD51*-like genes from *T. brucei* and *H. sapiens* were compared (Table 3.3). This comparison shows that the levels of sequence conservation between the RAD51 and DMC1(A) sequences is high (45-60% identity), but for the other RAD51-like proteins only low levels of sequence identity can be found (8-24%). On this basis, no clear orthologues can be discerned.

		H. sapiens							
		RAD51	DMC1	XRCC2	XRCC3	RAD51B	RAD51C	RAD51D	T.brucei RAD51
T. brucei	RAD51	60.1 70.8	44.8 57.6	14.3 21.7	20.3 32.0	20.2 33.6	19.5 33.3	16.2 25.5	100 100
	DMC1	52.7 67.3	57.0 70.1	14.6 24.8	20.7 33.0	23.8 38.2	19.7 34.1	14.8 26.6	49.2 62.6
	RAD51-4	15.5 29.4	15.1 27.2	10.9 18.9	17.4 28.9	15.8 28.1	15.7 25.9	16.4 27.5	15.6 29.1
	RAD51-3	17.7 27.6	18.1 26.0	12.4 19.4	16.3 27.8	17.0 26.4	22.8 33.5	10.8 20.3	17.0 25.9
	RAD51-6	16.8 23.7	15.7 25.7	8.0 15.9	17.0 26.4	14.9 25.1	15.6 27.0	11.7 20.1	15.9 26.3
	RAD51-5	11.3 22.8	9.5 22.2	13.5 23.0	11.8 21.1	12.1 23.7	12.2 24.3	13.5 22.7	10.8 25.2
	H.sapiens RAD51	100 100	51.6 65.1	16.1 24.9	21.0 33.6	24.7 37.1	22.2 33.1	15.8 27.0	

Table 3.2: The sequence homology of the RAD51-like proteins of *T. brucei* and *H. sapiens*. The sequence identities (large numbers) and similarities (small numbers) when comparing each of the *T. brucei* and *H. sapiens* RAD51-like proteins are shown. The identities and similarities were generated through pair wise comparison using AlignX (Vector NTI).

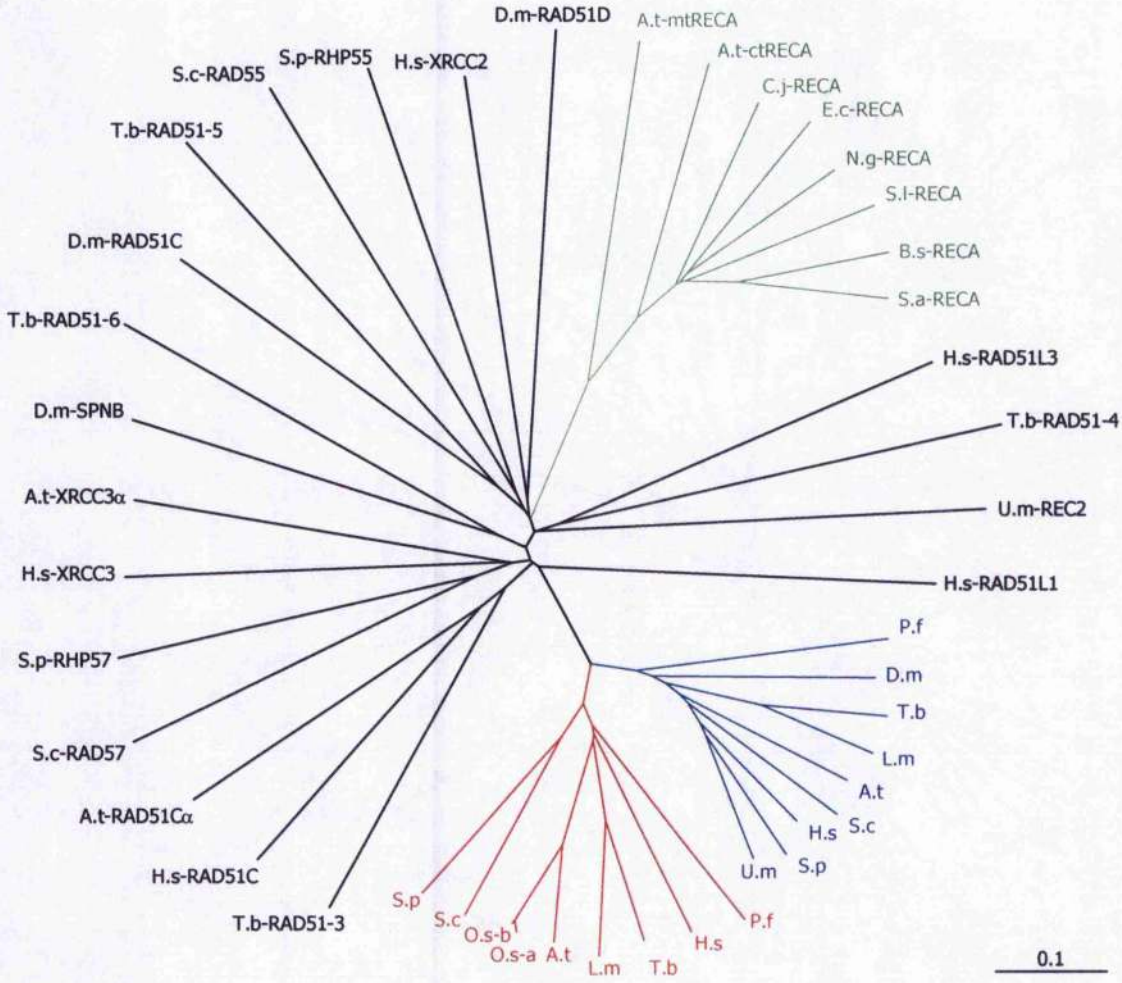


Figure 3.4: Phylogenetic tree of RecA and Rad51-like proteins. The polypeptide sequences of 9 Rad51, 9 Dmc1, 8 RecA and 19 Rad51-like genes from *T. brucei* (T.b-), *Leishmania major* (L.m-), *Plasmodium falciparum* (P.f-), *Saccharomyces cerevisiae* (S.c-), *Schizosaccharomyces pombe* (S.p-), *Homo sapiens* (H.s-), *Drosophila melanogaster* (D.m-), *Arabidopsis thaliana* (A.t-), *Oryza sativa* (O.s-), *Ustilago maydis* (U.m-). *Escherichia coli* (E.c-), *Campylobacter jejuni* (C.j-), *Neisseria gonorrhoeae* (N.g-), *Bacillus subtilis* (B.s-), *Streptomyces lividans* (S.l-) and *Staphylococcus aureus* (S.a-) were compared by ClustalW (Chenna *et al.*, 2003). The sequence comparison was then used to generate a phylogenetic tree and visualised using Treeview (Page, 1996). The Rad51 sequences are shown in blue, the Dmc1 sequences in red, the RecA sequences in green and the Rad51-like sequences in black.

It is interesting that the number of *Rad51*-like genes that an organism possesses varies. It may be that the overall number of *Rad51*-like genes is related to the size of the genome (Table 3.3), since multi-cellular eukaryotes such as *H. sapiens*, *A. thaliana* and *D. melanogaster* have 7, 6 and 6 *Rad51*-like genes, respectively, in total. In contrast, single-celled fungi such as *S. cerevisiae* and *S. pombe* contain 4 (although *S. pombe* may in fact contain a fifth Rad51-like gene; Grishchuk and Kohli, 2003). *T. brucei*, with 6 *RAD51*-like genes, is either unusual or the relationship between genome size and *Rad51*-like gene number is incorrect. To examine this we looked for *RAD51*-like genes in the genomes of the related Kinetoplastids, *Trypanasoma cruzi* and *Leishmania major*. To date, only the canonical *RAD51* has been detailed in *L. major* (McKean *et al.*, 2001). BLASTp searching the predicted genes in each genome, initially using T. brucei RAD51 and the canonical RAD51 from the two other parasites, identified 6 RAD51-related genes in *T. cruzi* and 5 in *L. major* (Table 3.3). Sequence comparisons of the predicted polypeptides within each genome are shown in Figures 3.5 and 3.6.

H. sapiens	RAD51	DMC1	5 RAD51-like
A. thaliana	RAD51	DMC1	4 RAD51-like
D. melanogaster	RAD51		4 RAD51-like
S. cerevisiae/pombe	RAD51	DMC1	2 RAD51-like
L. major	RAD51	DMC1	3 RAD51-like
T. brucei/T. cruzi	RAD51	DMC1	4 RAD51-like
	strand exchange mitosis/meiosis	strand exchange meiosis	strand exchange co-factors?

Table 3.3: *Rad51*-like gene numbers in eukaryotes. The number of *Rad51* genes found in a range of eukaryotes and their potential function are shown.

The polypeptide sequences of the RAD51-like proteins found in *T. cruzi* and *L. major* are also equally divergent in comparison to each other as the RAD51-like proteins of *T. brucei*, suggesting that this divergence is not a specific, unusual trait for *T. brucei*. However, comparison of the polypeptide sequences of the RAD51-like proteins found in *T. cruzi* (Table 3.5) and *L. major* (Table 3.6) to those found in *T. brucei* suggests that each of the *T. brucei* RAD51-like proteins appears to contain an orthologue in *T. cruzi* and *L. major*. The exception to this is *T. brucei* RAD51-5, where an orthologue cannot be found in *L. major*. Whether this is because the *L. major Rad51-5* is too diverged to be identified, has not been sequenced yet, or is truly absent is not known. This result

suggests that it is a trait of Kinetoplastids to contain an unusually large number of *RAD51*-related genes that show high levels of sequence divergence.

		T. cruzi					
		RAD51	DMC1	RAD51-3	RAD51-4	RAD51-5	RAD51-6
	RAD51	79.9	48.5	18.1	17.1	12.8	16.7
	DMC1	48.2	88.0	18.5	15.9	11.7	15.9
T. brucei	RAD51-3	16.1	17.1	49.8	14.6	12.5	12.7
	RAD51-4	14.6	18.2	14.3	39.9	15.6	13.3
	RAD51-5	11.4	12.1	12.4	12.4	34.8	13.8
	RAD51-6	15.3	17.2	16.4	13.8	17.1	39.2

Table 3.4: The sequence identities obtained when comparing the RAD51-like proteins from *T. brucei* and *T. cruzi*. The sequence identities obtained using AlignX (Vector NTI) to compare each of the *T.cruzi* RAD51-like proteins to those of *T. brucei* are shown. The highest sequence identity for each *T. cruzi* protein is highlighted in blue, and they have been named based on this.

		L. major					
		RAD51	DMC1	RAD51-3	RAD51-4	RAD51-5	RAD51-6
	RAD51	77.5	46.5	9.6	15.6	X	10.0
	DMC1	47.7	72.4	10.3	13.7	X	9.0
T. brucei	RAD51-3	16.3	17.2	23.9	13.2	X	11.4
	RAD51-4	15.5	17.6	12.7	18.0	X	12.6
	RAD51-5	9.8	13.6	7.6	12.9	X	8.7
	RAD51-6	17.2	17.8	15.1	13.8	X	21.9

Table 3.5: The sequence identities obtained when comparing the RAD51-like proteins from *T. brucei* and *L. major*. The sequence identities obtained using AlignX (Vector NTI) to compare each of the *L. major* RAD51-like proteins to those of *T. brucei* are shown. The highest sequence identity for each *L. major* protein is highlighted in blue, and they have been named based on this.

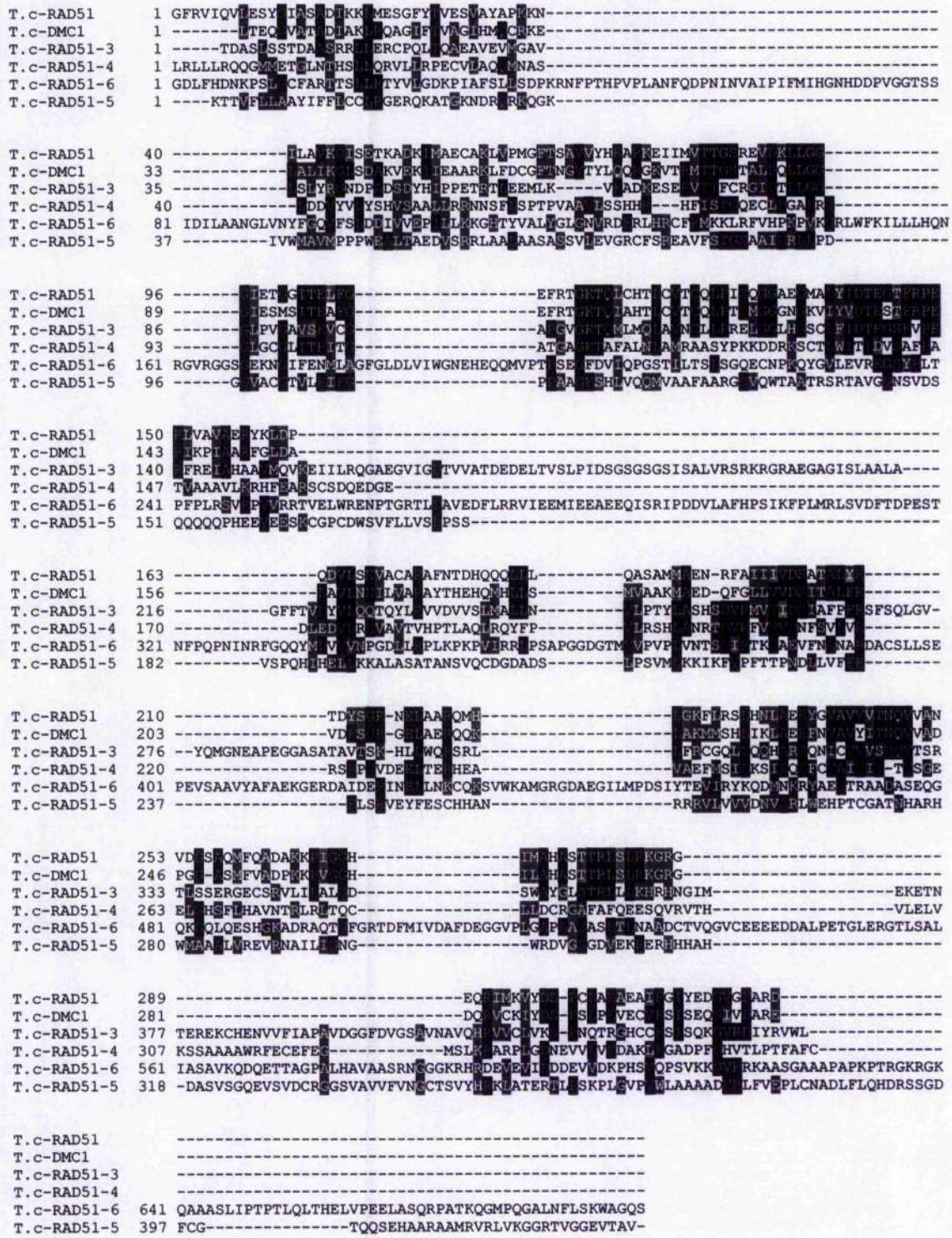


Figure 3.5: Alignment of the amino acid sequences of RAD51 and RAD51-like proteins of *T. cruzi*. *T. cruzi* (T.c-) RAD51 polypeptide sequence (top) is compared to those of the RAD51-like proteins found in *T. cruzi*. See appendix 2 for accession numbers. The alignment was carried out using ClustalW (Chenna *et al.*, 2003) and coloured using Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

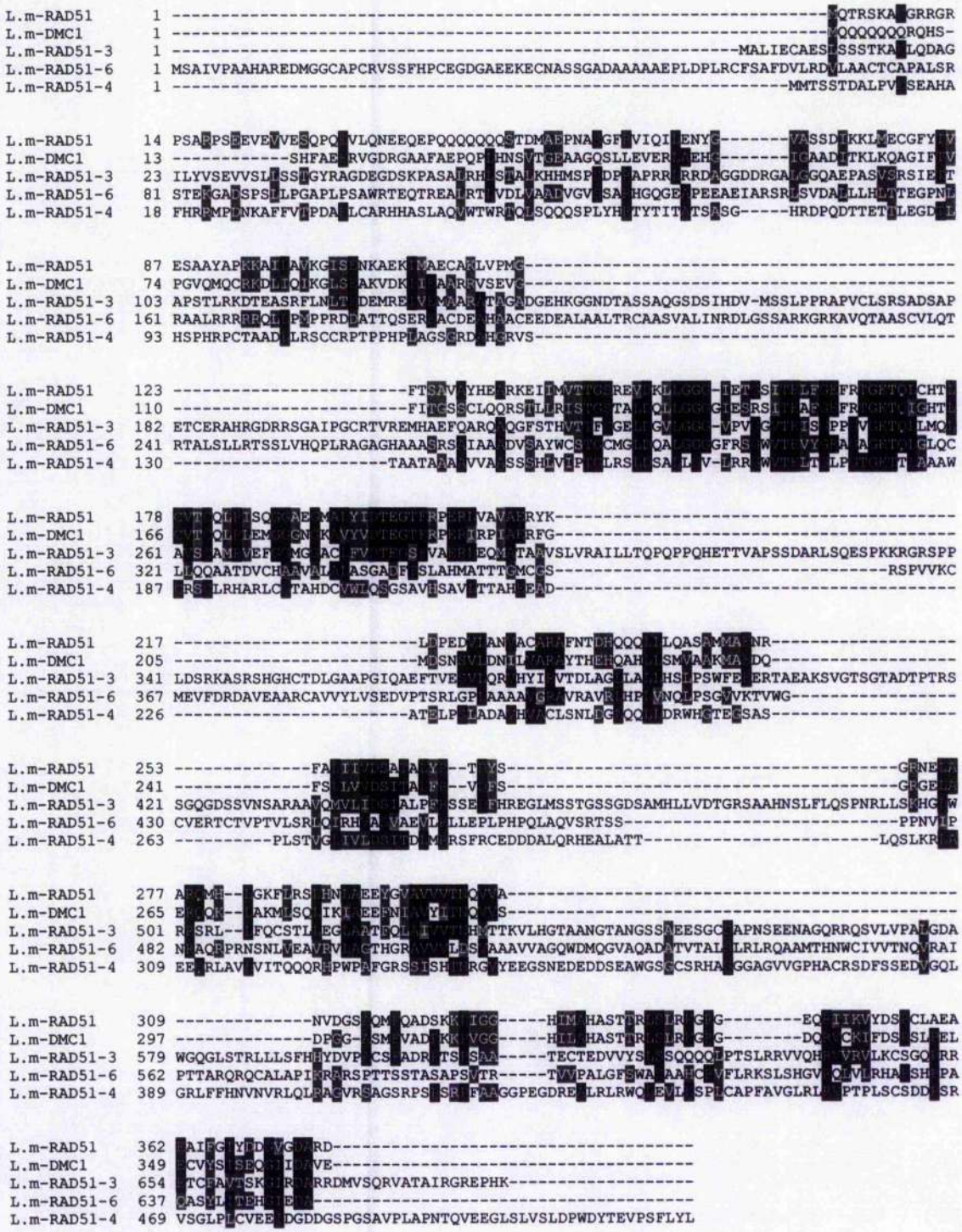


Figure 3.6: Alignment of the amino acid sequences of RAD51 and RAD51-like proteins of *L. major*. *L. major* (L.m-) RAD51 polypeptide sequence (top) is compared to those of the RAD51-like proteins found in *L. major*. See appendix 2 for accession numbers. The alignment was carried out using ClustalW (Chenna *et al.*, 2003) and coloured using Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

3.5 Sequence comparisons of *T. brucei* RAD51-3, RAD51-5 and DMC1

Functional characterisation of all of the *T. brucei RAD51*-like genes, by genetic manipulation, was deemed beyond the scope of this project, and three genes were therefore chosen for further analysis. These were *RAD51-3*, *RAD51-5* and *DMC1*. Further sequence analysis of these proteins and an explanation of their selection is provided below.

Sequence alignments of the predicted amino acid sequences of the *T. brucei* RAD51-3, RAD51-5 and DMC1 proteins were carried out to determine their homology to a number of putative orthologues, based on their best BLASTp hits. In these cases the ClustalW alignments were not manually manipulated to maximise homology. Alignment of the *T. brucei* RAD51 polypeptide sequence to Rad51 sequences from other organisms, carried out for comparison, showed 51-79% sequence identity (Fig 3.7). *T. brucei* RAD51-3 showed less homology to Rad51 proteins (in the range 12.2 to 17.7%) and was more homologous with Rad57 sequences from yeast and Rad51C from humans. Even here, the levels of sequence identity were low, however: between 13 and 33% (Fig 3.8). *T. brucei* RAD51-5, initially identified by searches with RecA, was first compared with bacterial RecA sequences. This showed low levels of homology, in the range of 9-29% sequence identity (Fig 3.9). However, comparison of *T. brucei* RAD51-5 with Rad51 sequences, including *T. brucei* RAD51, resulted in even lower identifiable homology: 8-21% sequence identity (Fig 3.10). This indicates that *T. brucei* RAD51-5 is substantially diverged, a result that is not unusual for a strand exchange enzyme (as has already been discussed). Very divergent strand exchange enzymes have also been described in *U. maydis*, which has been shown to encode two stand exchange enzymes: Rad51, which is highly homologous to other Rad51 proteins, and Rec2, which is much more highly divergent (Bennett and Holloman, 2001). Bacterial RecA has also been shown to operate in a eukaryotic system, suggesting that a strand exchange enzyme most homologous to a bacterial enzyme does not necessarily mean that the enzyme that will not function. *T. brucei* DMC1, like RAD51, showed high levels of sequence homology, with 53-73% sequence identity seen in comparison with Dmc1 sequences from other organisms (Fig 3.11). In comparison, it showed lower levels of sequence homology with Rad51 sequences (44 to 52.7% alignments not shown), suggesting that *T. brucei* DMC1 is most likely a counterpart of meiosis-specific Rad51 orthologues, although it should be noted that, of the sequences aligned in Fig 3.11, only the yeast, human and rice Dmc1 sequences have been shown to act during meiotic exchange.

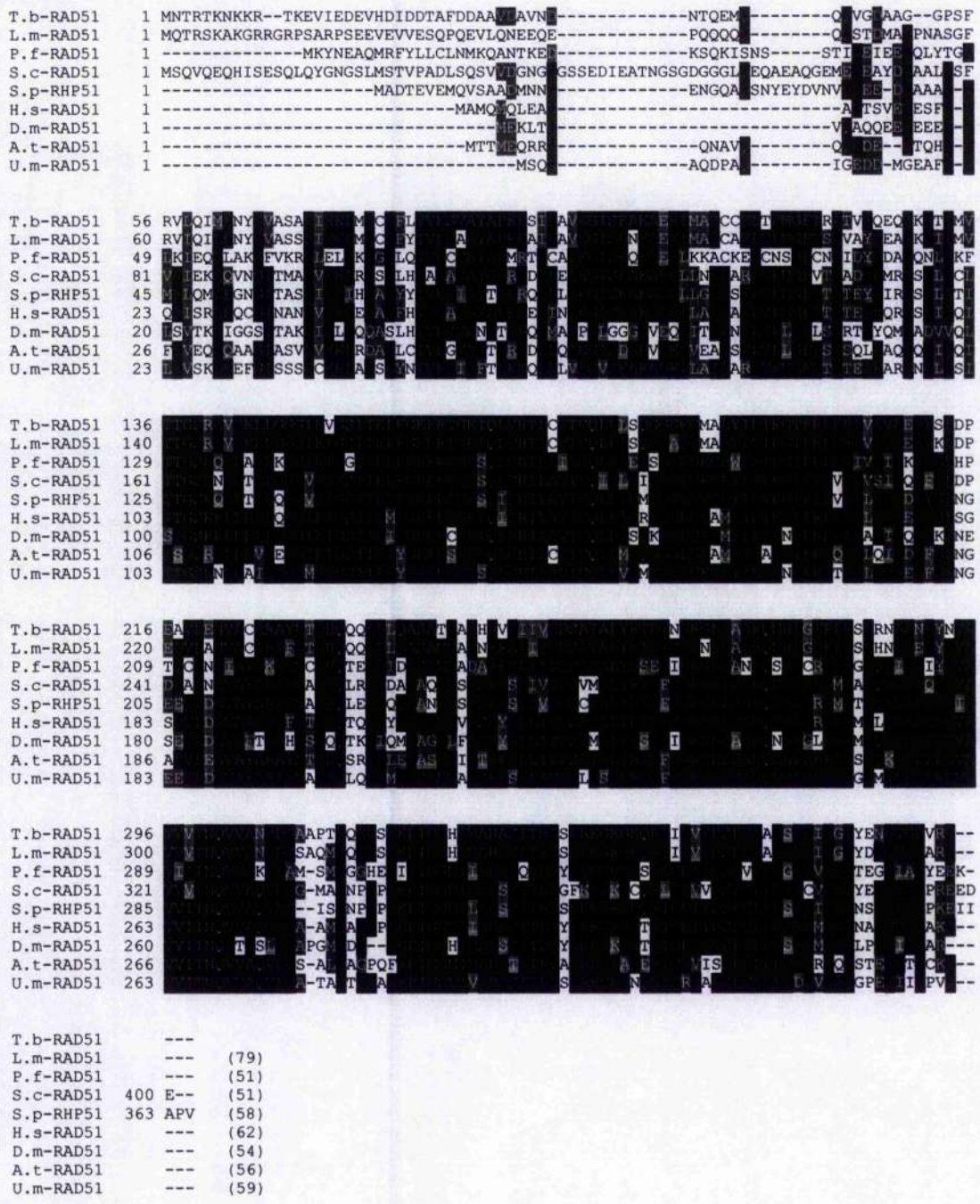


Figure 3.7: Alignment of the amino acid sequences of *T. brucei* RAD51 with other Rad51 proteins. The *T. brucei* (T.b-) RAD51 polypeptide sequence (top) is compared to the Rad51 sequences of *Leishmania major* (L.m-), *Plasmodium falciparum* (P.f-), *Saccharomyces cerevisiae* (S.c-), *Schizosaccharomyces pombe* (S.p-), *Homo sapiens* (H.s-), *Drosophila melanogaster* (D.m-), *Arabidopsis thaliana* (A.t-) and *Ustilago maydis* (U.m-). See appendix 2 for accession numbers. The identity values, calculated using AlignX, are shown (in brackets) for each of the Rad51 polypeptides compared to *T. brucei* RAD51. The alignment was carried out using ClustalW (Chenna *et al.*, 2003) and coloured using Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

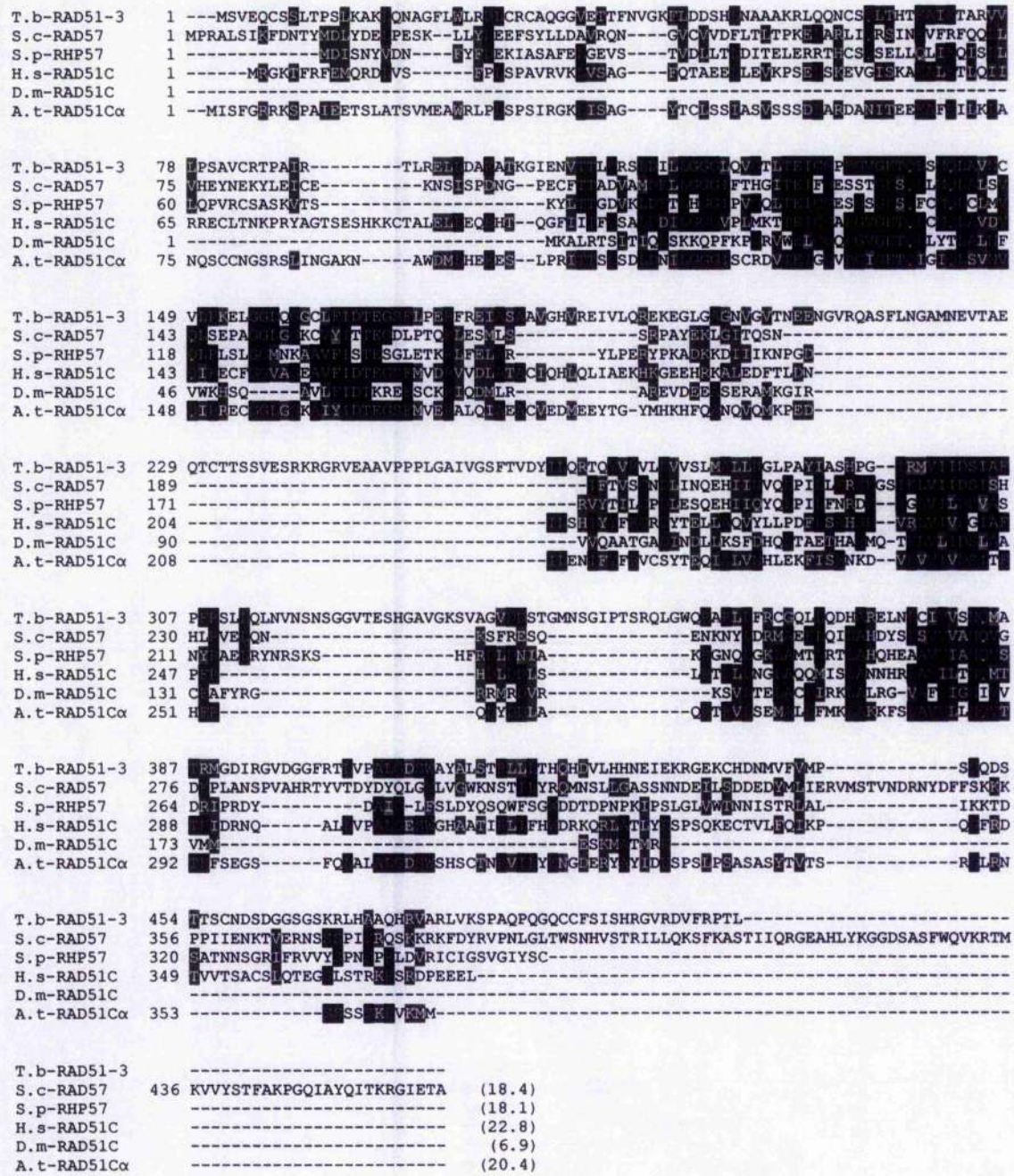


Figure 3.8: Alignment of the amino acid sequences of *T. brucei* RAD51-3 with Rad57 proteins. The *T. brucei* (T.b-) RAD51-3 polypeptide sequence (top) is compared to the Rad57 sequences of *Saccharomyces cerevisiae* (S.c-), *Schizosaccharomyces pombe* (S.p-), *Homo sapiens* (H.s-), *Drosophila melanogaster* (D.m-) and *Arabidopsis thaliana* (A.t-). See appendix 2 for accession numbers. The identity values, calculated using AlignX, are shown (in brackets) for each of the Rad57 polypeptides compared to *T. brucei* RAD51-3. The alignment was carried out using ClustalW (Chenna *et al.*, 2003) and coloured using Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

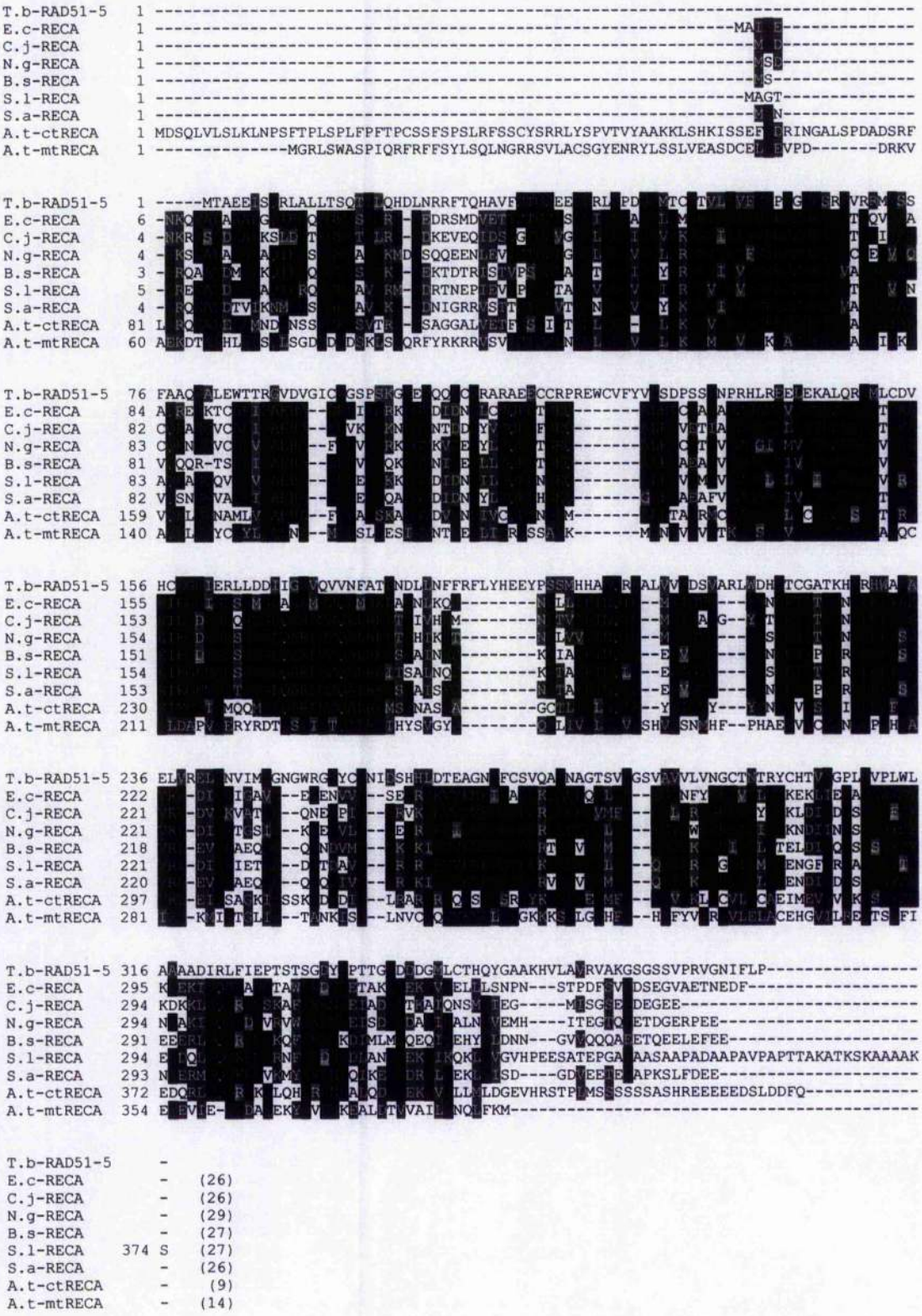


Figure 3.9: Alignment of the amino acid sequences of *T. brucei* RAD51-5 with RecA proteins. The *T. brucei* (T.b-) RAD51-5 polypeptide sequence (top) is compared to the RecA sequences of *Escherichia coli* (E.c-), *Campylobacter jejuni* (C.j-), *Neisseria gonorrhoeae* (N.g-), *Bacillus subtilis* (B.s-), *Streptomyces lividans* (S.l-), *Staphylococcus aureus* (S.a-) and two *Arabidopsis thaliana* RecA sequences with chloroplast (A.t-ct) and mitochondrial (A.t-mt) localisation. See appendix 2 for accession numbers. The identity values, calculated using AlignX, are shown (in brackets) for each of the RecA polypeptides compared to *T. brucei* RAD51-5. The alignment was carried out using ClustalW (Chenna *et al.*, 2003) and coloured using Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

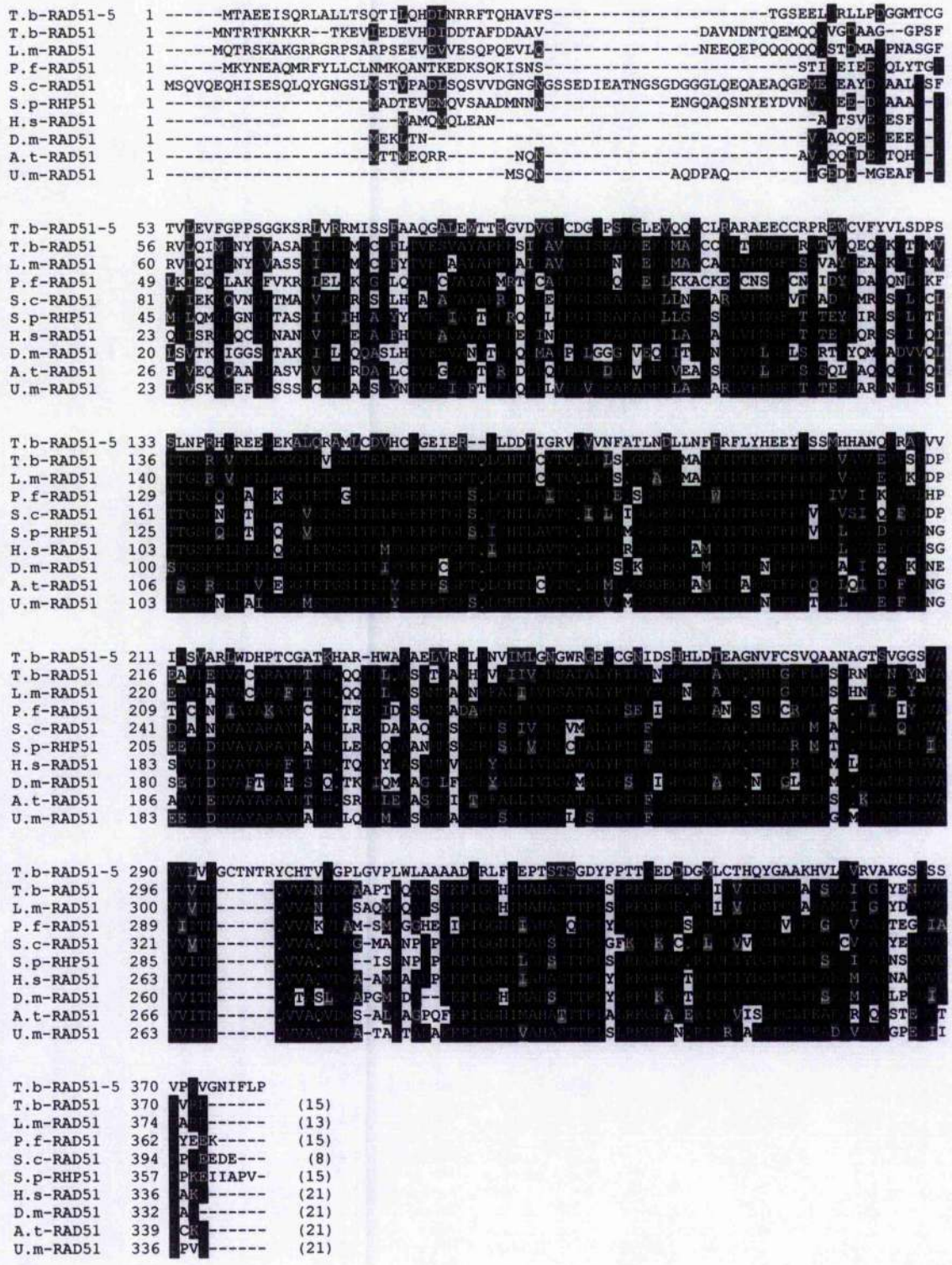


Figure 3.10: Alignment of the amino acid sequences of *T. brucei* RAD51-5 with Rad51 proteins. The *T. brucei* (T.b-) RAD51-5 polypeptide sequence (top) is compared to the Rad51 sequences as shown in figure 3.7. The identity values, calculated using AlignX, are shown (in brackets) for each of the Rad51 polypeptides compared to *T. brucei* RAD51-5. The alignment was carried out using ClustalW (Chenna *et al.*, 2003) and coloured using Boxshade 3.21 (http://www.ch.embnet.org/software /BOX_form.html).

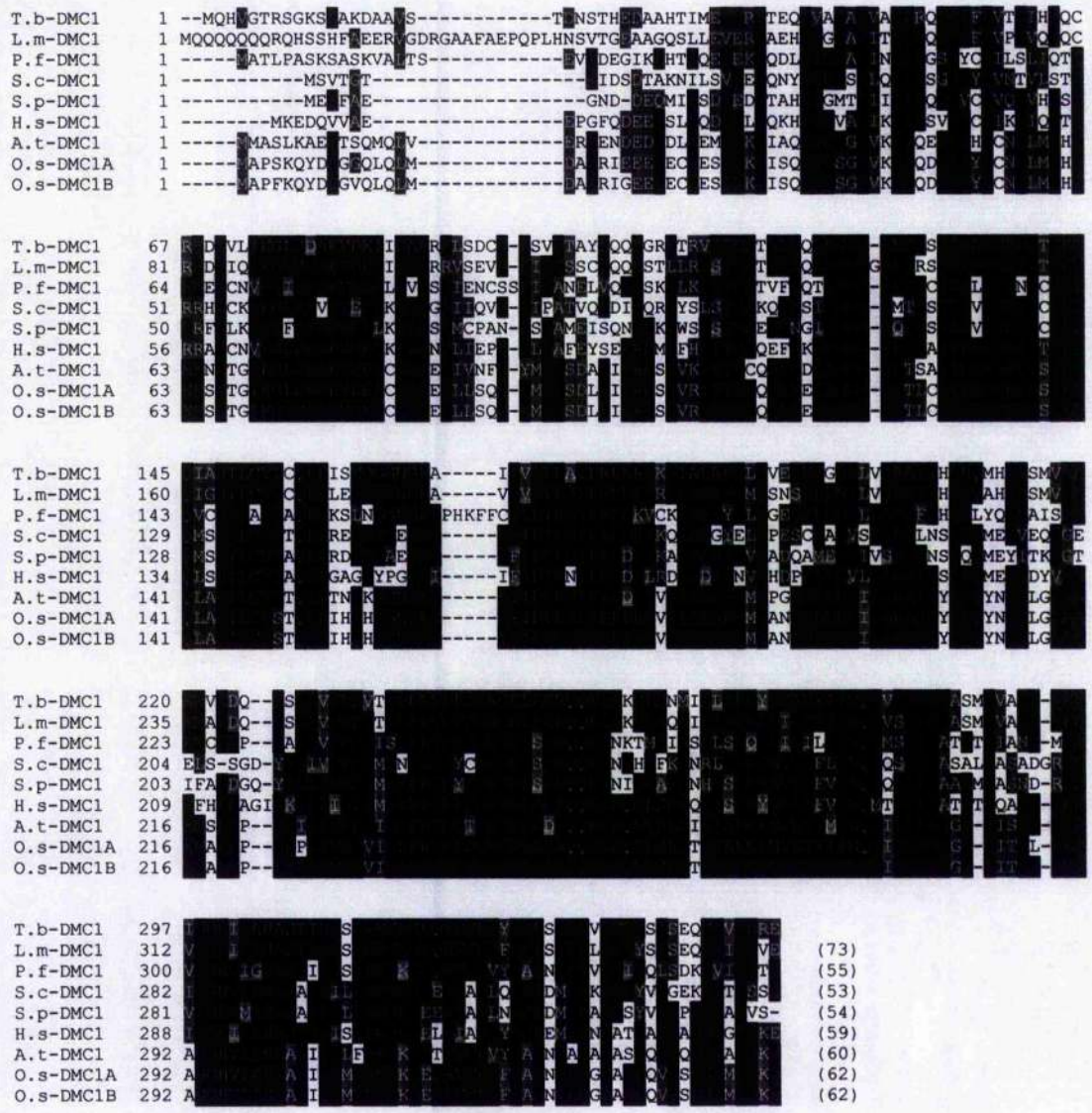


Figure 3.11: Alignment of the amino acid sequences of *T. brucei* DMC1 with other Dmc1 proteins. The *T. brucei* (T.b-) DMC1 polypeptide sequence (top) is compared to the Dmc1 sequences of *Leishmania major* (L.m-), *Plasmodium falciparum* (P.f-), *Saccharomyces cerevisiae* (S.c-), *Schizosaccharomyces pombe* (S.p-), *Homo sapiens* (H.s-), *Arabidopsis thaliana* (A.t-), *Drosophila melanogaster* (D.m-) and the two DMC1 sequences from *Oryza sativa* (O.s-). See appendix 2 for accession numbers. The identity values, calculated using AlignX, are shown (in brackets) for each of the Dmc1 polypeptides compared to *T. brucei* DMC1. The alignment was carried out using ClustalW (Chenna *et al.*, 2003) and coloured using Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

The remainder of this thesis considers the functions of *T. brucei RAD51-3*, *RAD51-5* and *DMC1* in bloodstream stage cells. These were chosen, in part, because they were the first *RAD51*-like genes to emerge intact from the genome sequencing initiative. However, they are interesting choices for a number of reasons. As has been argued above, it is not possible to infer their likely roles through sequence comparisons, with the possible exception of *DMC1*, which groups closely with meiosis-related RAD51 proteins and might therefore act in that process (Bingle *et al.*, 2001; Gibson, 2001; Gibson and Stevens, 1999). Each has the capacity to independently catalyse DNA strand exchange, rather than act as cofactors, given the retention of Walker A and B boxes. It would be interesting if *DMC1* had adopted roles in mitotic recombination, whilst *RAD51-3* is, arguably, the best conserved of the *RAD51*-like genes, beyond *RAD51* itself and *DMC1*. *RAD51-5* is the most diverged in sequence and may be absent in *L. major* (potentially indicating a function specific to *T. brucei* and *T. cruzi*).

3.6 Analysis of copy number of *T. brucei RAD51-3*, *RAD51-5* and *DMC1*

To determine the number of copies of *RAD51-3*, *RAD51-5* and *DMC1* in the *T. brucei* genome, Southern analysis was carried out. Genomic DNA from 427 bloodstream form cells was restriction digested using *Bam*HI, *Eco*RI, *Eco*RV, *Hin*dIII and *Sac*II, separated by gel electrophoresis on a 0.8% 1 x TBE gel and transferred to a nylon membrane by Southern blotting. The blots were then probed with the entire open reading frame of each gene, which were PCR–amplified using Taq DNA polymerase and primers 'For' and 'Rev' for each gene (Appendix 1). *RAD51*, which had been shown previously to be single copy (McCulloch and Barry, 1999) was examined also as a control. The individual digestions, probed with each open reading frame, in general resulted in the production of a single band. The exceptions to this were *Eco*RV for *RAD51*, *Bam*HI for *RAD51-3*, *Eco*RV for *RAD51-5* and *Sac*II for *DMC1*; in each case the restriction enzyme cuts within the open reading frame, resulting in the production of two bands (Fig 3.12). The *Bam*HI digestion for *DMC1* should also have produced two bands in this analysis, as it cuts in the open reading frame, but only a single band is visible. This is most readily explained by a base pair difference between the 927 strain sequenced for the *T. brucei* genome database and the 427 strain utilised during this assay. Despite this, the result suggests that the *RAD51-3*, *RAD51-5* and *DMC1* are single copy.

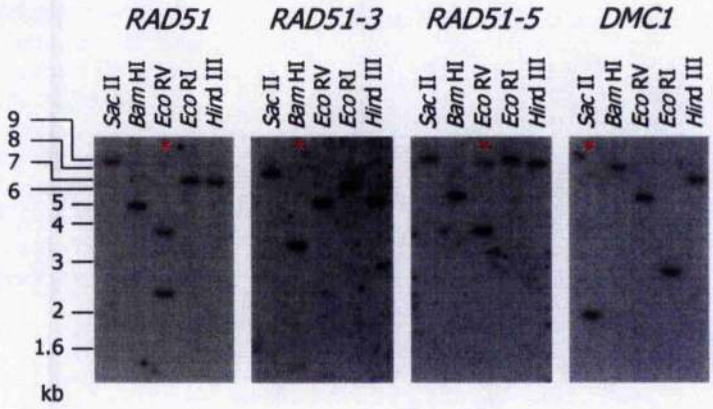


Figure 3.12: Southern analysis of the copy number of the *RAD51*-like genes. Genomic DNA from 427 wild type cells was digested with a range of enzymes (indicated) and probed with the open reading frame of one of the *RAD51*-like genes. Lanes marked with an asterisk (*) are those where two bands are visible.

3.7 Analysis of life cycle stage expression of *RAD51-3*, *RAD51-5* and *DMC1*.

During the life cycle of *T. brucei* the expression levels of many genes differ depending on what stage the cell is in. Examples of this include the *VSG*, which is expressed only in the bloodstream, and procyclin, which is expressed only in the procyclic stage (Pays *et al.*, 1994). The stages at which the control of these expression changes have been shown to occur encompass transcription initiation, mRNA processing, mRNA stability and protein production (Campbell *et al.*, 2003; D'Orso *et al.*, 2003; Vanhamme *et al.*, 1995). To examine if any of the *RAD51*-like genes are expressed in a lifecycle stage-dependant manner, the levels of their mRNA was assayed. *RAD51* was analysed also as a form of control, as it has already been shown to function in DNA repair and recombination in bloodstream and procyclic form cells (McCulloch and Barry, 1999). This question is particularly relevant for *DMC1*, as it has been shown to be most homologous to meiosis-specific *RAD51* homologues, DMC1, in various organisms (Bishop *et al.*, 1992; Doutriaux *et al.*, 1998; Fukushima *et al.*, 2000; Masson *et al.*, 1999; Staeva-Vieira *et al.*, 2003). If *DMC1* is truly a meiosis-specific factor, it is possible that its expression could be limited to the *T. brucei* salivary gland stage, where meiosis is proposed to take place in *T. brucei* (Bingle *et al.*, 2001). To perform this analysis, RT-PCR was carried out on cDNA (P. Blundell, gift) generated from total RNA from five life cycle stages: EATRO 795 pleomorphic bloodstream form cells, tsetse fly midgut-derived cells, salivary gland-derived cells, early bloodstream form cells (these are tsetse fly transmitted *T. brucei* that have been injected into mice and continue to express Metacyclic VSG; Barry and McCulloch, 2001) and Lister 427 monomorphic bloodstream form cells. In each case cDNA samples prepared without reverse transcriptase were also used to control for genomic DNA contamination. The RT-PCR was carried out using oligonucleotide primers specific for each of the genes.

For *RAD51*, the primers *RAD51*-J1 and *RAD51*-U3 (Appendix 1) were used, which should give a product of 401 bp (Fig 3.13). For *RAD51-3*, the primers *RAD51-3*-D1 and *RAD51-3*-U1 (Appendix 1) were used, which should give a product of 586 bp (Fig 3.13). For *RAD51-5*, the primers *RAD51-5*- D2 and *RAD51-5*-U1 (Appendix 1) were used, which should give a product of 178 bp (Fig 3.13). Finally, for *DMC1*, the primers *DMC1*-For and *DMC1*-U1 (Appendix 1) were used, which should give a product of 402 bp (Fig 3.13). In each case, a specific product of the appropriate size was generated in each of the life cycle stages tested, with no genomic DNA contamination in the reverse transcriptase minus control.

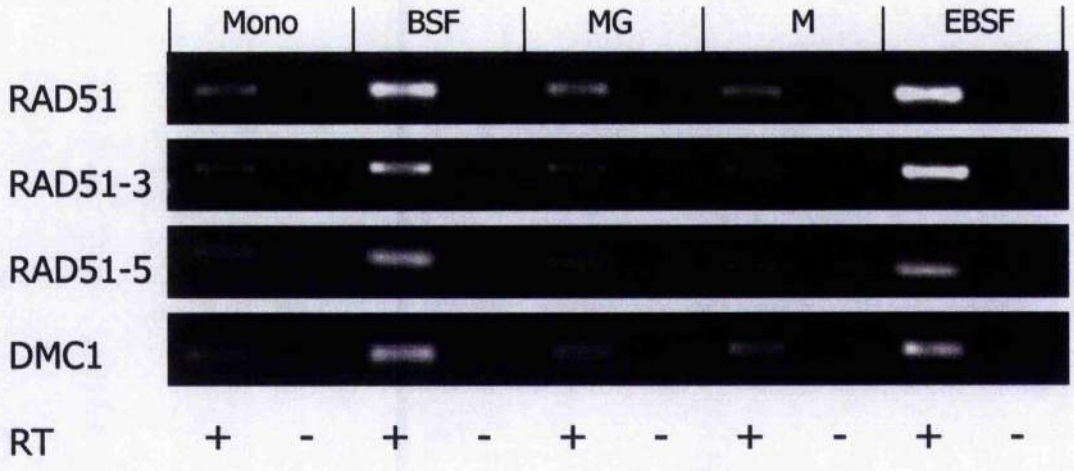


Figure 3.13: The expression of the *RAD51*-like genes in different life cycle stages. RT-PCR was carried out on cDNA generated from total RNA from trypanosomes from five life cycle stages. Gene specific primers were used to assay expression of the genes in each of the life cycle stages using reverse transcriptase plus (RT+) and minus (RT-) cDNA samples. Mono: monomorphic Lister 427 bloodstream form. BSF: EATRO 795 bloodstream form. MG: mid-gut derived *T. brucei*. M: metacyclic derived *T. brucei*. EBSF: early BSF EATRO 795 *T. brucei*.

This result suggests that each of the *RAD51*-like genes, including *DMC1*, are expressed, or at least transcribed, continually throughout the life cycle. Given the limited number of *T. brucei* cells recoverable during tsetse fly transmission, there was insufficient cDNA to allow more quantitative PCR approaches to be used which might have indicated changes in mRNA abundance of the different genes. As specific antibodies directed against the different proteins are not available, this could not be pursued further by Western analysis.

3.8 Analysis of *RAD51*-like mRNA level in response to DNA damage

Many organisms have been shown to increase expression of repair genes in response to DNA damage. One of the best characterised is the SOS response of *E. coli*, and other bacteria, involving the LexA repressor that, in the absence of DNA damage, prevents expression of the SOS genes (Kuzminov, 1999). In response to DNA damage bacterial RecA is activated when it forms a nucleoprotein filament on single strand DNA formed as a result of the cells attempt to replicate damaged DNA (Smith and Walker, 1998). Active RecA then induces self-cleavage of the LexA repressor, leading to the induced expression of approximately 20 repair genes (Kuzminov, 1999). One of the genes that becomes up-regulated is RecA itself, the level of which can increase as much as 50-fold. The same SOS response is not present in the archae or in eukaryotes, but here also RadA and Rad51 up-regulation is seen. For example, in *S. cerevisiae* and *S. pombe* Rad51 was shown to be induced (approximately 3 to 5-fold) in response to DNA damage, along with 25 other genes (Essers *et al.*, 2002; Lisby *et al.*, 2004; Mercier *et al.*, 2001). In *Tetrahymena thermophillia*, Rad51 levels have been suggested to increase by up to 100-fold in response to DNA damage (Campbell and Romero, 1998). Similar observations have also been made in protozoa, as both *L. major* (McKean *et al.*, 2001) and *P. falciparum* (Bhattacharyya and Kumar, 2003) were observed to increase *Rad51* expression in response to DNA damage. Exceptions to this general rule do exist, however. There is little evidence for increased Rad51 in mammalian cells exposed to DNA damage. Moreover, some archae induce RadA expression (Reich *et al.*, 2001), whilst others do not (Komori *et al.*, 2000) and some bacteria appear to lack an SOS response (Black *et al.*, 1998).

In light of this knowledge, we wanted to test whether or not *T. brucei* operates in a similar manner and up-regulates the expression of *RAD51*, or any of the three *RAD51*-like genes, in response to DNA damage. The potential DNA damage induction of *RAD51*-like genes has not been widely studied, but XRCC3 and Rad51C in *Arabidopsis thaliana* are modestly induced following γ irradiation (Osakabe *et al.*, 2002). For *T. brucei* we could analyse only the mRNA level, as antibodies for the RAD51-like proteins were not available. However, if it emerged that changes in mRNA levels, whether through altered transcription rates or changes in mRNA stability, were a general feature of the putative repair genes, it may provide a quick, simple assay to define any genes involved in DNA damage repair.

3.8.1 Analysis of the mRNA level of *RAD51* and the 3 *RAD51*-like genes

From the previous RT-PCR, we already know that each of the genes is expressed, or at least transcribed, in bloodstream form cells (section 3.4). However, this analysis was conducted in a non-quantitative manner and suggests little about the relative expression level of each of the genes. Therefore, before the analysis of the expression of the *RAD51*-like genes following DNA damage could be performed, the quantity of RNA required to produce a suitable signal on a Northern blot had to be assessed. To do this, total RNA was extracted (RNeasy Mini Kit, Qiagen) from 100 mls of bloodstream form Lister 427 wild type cells, grown to a density of 2 x 10^6 cells.ml^{-1}, and quantified by spectrophotometer (Beckman DU650 spectrophotometer). Four repetitions of 5, 10 and 20 µg samples were separated by electrophoresis on a denaturing formaldehyde gel (section 2.10.1) before being transferred to a nylon membrane by Northern blotting (section 2.10.2). The blots were then probed with the entire open reading frame of one of the 3 *RAD51*-like genes, or with *RAD51*. In each case a single specific band was generated in each lane which we assumed to be mature mRNA (Fig 3.14). All four of the genes produced a detectable signal with 5 µg of total RNA, including *DMC1*. This indicates that each gene produces a putative mRNA in bloodstream form cells, although it is not possible to assess if they are all translated to protein, nor the relative stability of the RNA species.

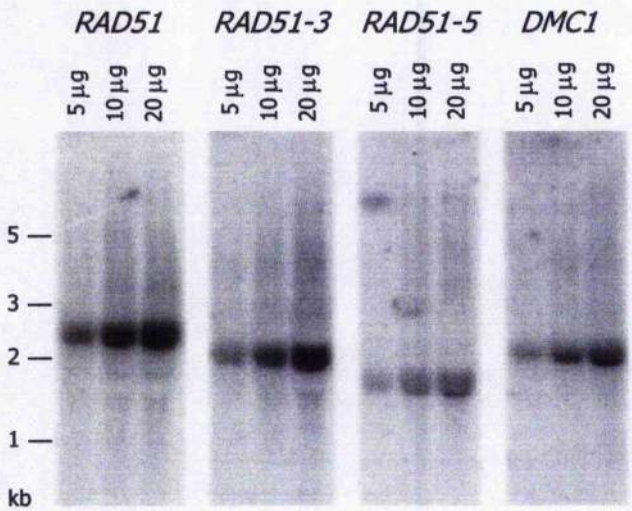


Figure 3.14: Northern analysis of *RAD51* and the three *RAD51*-like genes. The quantity of total RNA loaded in each lane is indicated and size markers are shown.

3.8.2 Analysis of the effect of phleomycin

The DNA damaging agent chosen for use during the analysis of potential changes in mRNA levels was phleomycin. Two primary reasons dictated this choice. First, it is a compound known to directly cause double strand breaks (Giloni *et al.*, 1981), in contrast to other DNA damage agents often used, such as Methylmethane Sulphonate (MMS; Sedgwick, 2004). Second, it was the DNA damaging agent used during the analysis of *RAD51* expression in the related parasite *L. major* (McKean *et al.*, 2001). Before examining the effect on *T. brucei RAD51* expression, an attempt was made to examine the effect of phleomycin on *T. brucei* genomic DNA and on growth. To do this, bloodstream form Lister 427 wild type cells were grown to a density of 2 x 10^6 cells.ml^{-1}, then diluted to 1.5 x 10^6 cells.ml^{-1} and phleomycin was added to 10 μg.ml^{-1}. From this, cell counts were made for a period of 24 hours and 5 x 10^7 cells were used to generate genomic plugs (section 2.9.1) at 0, 4, 8 and 24 hour time points. The growth following DNA damage is shown in Fig 3.15, from this it is apparent that over 24 hours the concentration of phleomycin causes a cessation of population doubling, but little decrease in cell number.

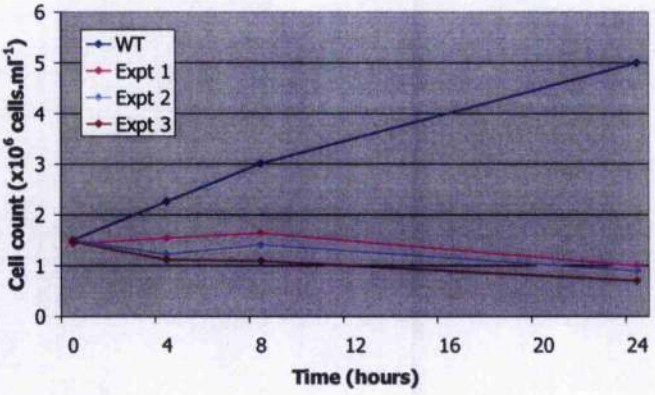


Figure 3.15: Growth effect of 10 μg.ml^{-1} phleomycin on Lister 427 wild type cells. The growth counts carried out during exposure to 10 μg.ml^{-1} phleomycin for three experiments (EXPT 1,2 and 3) are shown. The result for a control experiment (WT), where cells were not exposed to phleomycin, is also shown.

To examine the effect of phleomycin on the *T. brucei* genome, the genomic plugs were placed in the wells of a 1% agarose gel (Seakem LE) and the DNA was separated by pulsed field electrophoresis (section 2.10). The gel was then Southern blotted before being probed with β-tubulin (Fig 3.16). From this result, a faint smear was visible below the intact chromosomal band (>194 kbp), which appeared to increase in intensity as time increased. This smear is most likely the result of the phleomycin causing DNA breaks, resulting in fragmented chromosomes that are able to migrate further into the gel during electrophoresis. The increase in intensity of the smear is due to the increase in

exposure to the DNA damaging agent. Although it is possible that damage may have occurred during the preparation of the genomic plugs rather than by phleomycin-treatment, or that the later time points have more DNA loaded, no convenient control is available, as all *T. brucei* genomic DNA is likely to be affected by phleomycin damage. The approximately equal intensity of the intact chromosomal bands suggests that the loading of the samples is approximately equal. However, when looking at the chromosomal band in the ethidium gel it may appear as if the 24 hour time point has had more DNA loaded compared to the other time points. This is probably the result of damage to the chromosomes resulting in a more disperse, overexposed band. Even if the 24 hour time point were to have more DNA loaded it still appears that damage is being caused as after 8 hours a smear, which is under 49 kbp in size, is visible that is not present at 0 and 4 hour time points. Overall, the results of this assay suggest that phleomycin has caused DNA breaks in this assay.

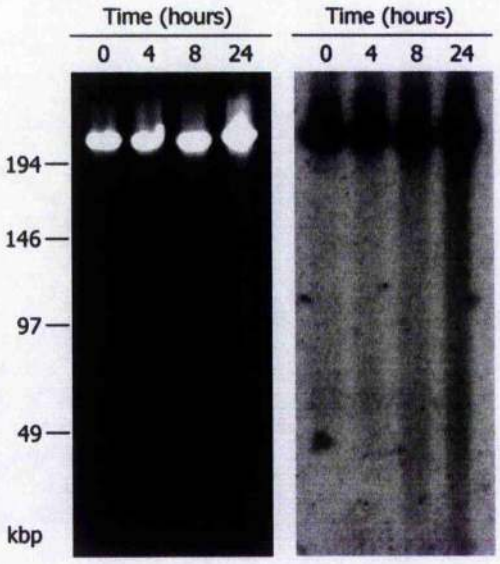


Figure 3.16: PFGE of Lister 427 wild type cells exposed to 10 μg.ml^{-1} phleomycin. Lister 427 wild type cells were grown in the presence of 10 μg.ml^{-1} phleomycin for 0, 4, 8 and 24 hours. The DNA was then separated by PFGE and subsequently the gel was Southern blotted and probed with β-tubulin. Size markers (Lambda CHEF DNA ladder, BioRad) are indicated.

3.8.3 Analysis of the effect of phleomycin on *RAD51*, *RAD51-3*, *RAD51-5* and *DMC1* mRNA levels

The mRNA levels of *RAD51* and the 3 *RAD51*-like genes were analysed at 0, 4, 8 and 24 hour time points after exposure to 1, 2 or 10 μg.ml^{-1} phleomycin. Three repetitions were carried out at 10 μg.ml^{-1} and 2 repetitions at 1 and 2 μg.ml^{-1}. Bloodstream form Lister 427 wild type cells were grown to a density of 2 x 10^6 cells.ml^{-1} before being diluted to 1.5 x 10^6 cells.ml^{-1} containing phleomycin. Total RNA was then extracted

from 100 mls of cells at 0, 4, 8 and 24 hour time points and quantified by spectrophotometer (Beckman DU650 spectrophotometer). Four repetitions of 5 µg samples from each time point were separated by electrophoresis on a denaturing formaldehyde gel and subsequently transferred to a nylon membrane by Northern blotting. The blots were then probed with the open reading frame of one of the 3 *RAD51*-like genes, or *RAD51*. The bands present in each of the developed blots were quantified (Image Quant) before the blots were stripped (section 2.11.3), re-probed with *GPI-8* (Lillico *et al.*, 2003) and re-quantified (Fig 3.17).

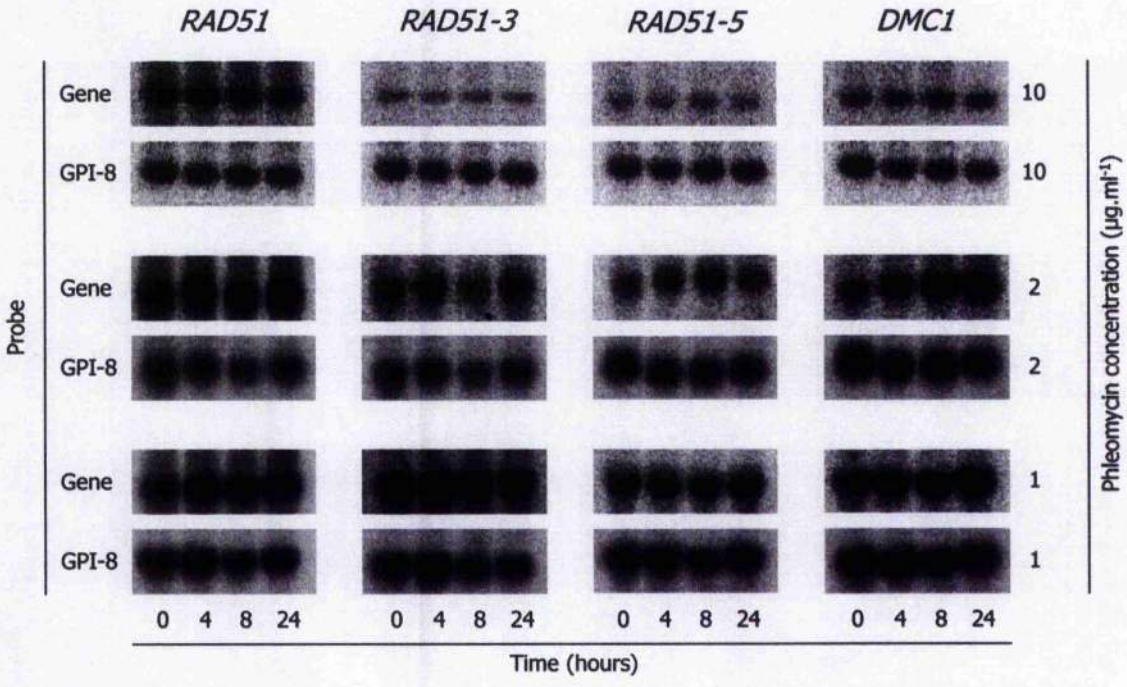


Figure 3.17: Northern analysis of the mRNA levels of *RAD51* and the 3 *RAD51*-like genes in response to phleomycin. Examples of the Northern blots of total RNA probed with the open reading frame of the gene of interest (Gene) which were stripped and re-probed with GPI-8 are shown. Cells were exposed to varying concentrations of phleomycin for 0, 4, 8 or 24 hours.

GPI-8 was used as a single copy gene control as its expression level should not alter in response to DNA damage, since it is involved in surface molecule biosynthesis (Lillico *et al.*, 2003). For both the *RAD51*-related genes and *GPI-8* values determined at 0 hours were counted as 100% and the values for the 4, 8 and 24 hour time points calculated relative to 100%. By then comparing the *RAD51* and *RAD51*-like gene values to the *GPI-8* values, any changes in mRNA levels in response to DNA damage could be determined, whilst removing any error due to loading. The results of this analysis are shown in Tables 3.7 and 3.8 and graphed in figure 3.18.

A

Probe		GENE				GPI-8			
	Time (hours)	Expt 1	Expt 2	Avge	St dev	Expt 1	Expt 2	Avge	St dev
RAD51	0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	0.0
	4	112.8	93.2	103.0	13.9	106.1	74.2	90.2	22.6
	8	117.5	128.3	122.9	7.6	79.8	71.0	75.4	6.2
	24	149.6	134.7	142.2	10.5	94.7	78.0	86.4	11.8
RAD51-3	0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	0.0
	4	103.7	75.0	89.4	20.3	98.9	75.4	87.2	16.6
	8	99.8	91.6	95.7	5.8	66.7	65.2	66.0	1.1
	24	122.9	106.7	114.8	11.5	84.0	71.0	77.5	9.2
RAD51-5	0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	0.0
	4	92.6	87.0	89.8	4.0	109	76.9	93.0	22.7
	8	90.0	108.6	99.3	13.2	77.1	64.0	70.6	9.3
	24	103.3	103.6	103.5	0.2	94.1	79.4	86.8	10.4
DMC1	0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	0.0
	4	100.7	74.4	87.6	18.6	92.2	77.4	84.8	10.5
	8	111.4	107.8	109.6	2.5	75.2	67.8	71.5	5.2
	24	163.8	148.8	156.3	10.6	87.7	85.0	86.4	1.9

B

Probe		GENE				GPI-8			
	Time (hours)	Expt 1	Expt 2	Avge	St dev	Expt 1	Expt 2	Avge	St dev
RAD51	0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	0.0
	4	99.8	119.3	109.6	13.8	78.4	110.9	94.7	23.0
	8	98.0	133.0	115.5	24.7	64.8	101.3	83.1	25.8
	24	132.2	144.6	138.4	8.8	81.8	111.8	96.8	21.2
RAD51-3	0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	0.0
	4	99.5	110.0	104.8	7.4	88.8	108.2	98.5	13.7
	8	79.2	116.5	97.9	26.4	70.0	114.3	92.2	31.3
	24	112.2	126.3	119.3	10.0	99.1	127.3	113.2	19.9
RAD51-5	0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	0.0
	4	94.8	111.1	103.0	11.5	109.3	87.7	98.5	15.3
	8	82.7	117.8	100.3	24.8	99.7	69.5	84.6	21.4
	24	107.7	106.2	107.0	1.1	101.2	101.6	101.4	0.3
DMC1	0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	0.0
	4	109.7	125.5	117.6	11.2	93.1	107.8	100.4	10.4
	8	127.0	167.7	147.4	28.8	79.1	102.8	91.0	16.8
	24	171.4	193.6	182.5	15.7	108.7	115.7	112.2	4.9

Table 3.6: The mRNA levels of *RAD51* and the 3 *RAD51*-like genes in response to 1 and 2 μg.ml^{-1} phleomycin. The results for the two repetitions of the DNA damage experiments (Expt 1 and 2) carried out at 1 (A) and 2 (B) μg.ml-1 phleomycin are shown. The values represent the quantification of Northern blots probed with *RAD51*, *RAD51-3*, *RAD51-5* or *DMC1* (GENE) and their subsequent quantification after re-probing with *GPI-8* (GPI-8): the quantification value obtained for time 0 hours was regarded as 100% and all other values scaled accordingly. Averages (Avge) and the standard deviations (St dev) generated from the data are also shown.

	Probe	GENE					GPI-8				
	Time (hours)	Expt 1	Expt 2	Expt 3	Avge	St dev	Expt 1	Expt 2	Expt 3	Avge	St dev
RAD51	0	100.0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	100.0	0.0
	4	158.4	93.4	91.3	114.4	38.1	131.3	93.2	87.8	104.1	23.7
	8	133.7	88.2	87.3	103.1	26.5	120.1	96.0	95.0	103.7	14.2
	24	137.5	112.7	84.0	111.4	26.8	117.4	122.6	99.2	113.1	12.3
RAD51-3	0	100.0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	100.0	0.0
	4	83.8	87.2	81.9	84.3	2.7	89.2	87.8	86.0	87.7	1.6
	8	80.3	92.4	84.1	85.6	6.2	86.9	94.9	92.6	91.5	4.1
	24	97.3	100.5	82.3	93.4	9.7	104.9	124.8	99.5	109.7	13.3
RAD51-5	0	100.0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	100.0	0.0
	4	89.0	100.2	92.8	94.0	5.7	81.3	97.4	81.9	86.9	9.1
	8	88.0	97.2	100.4	95.2	6.4	77.7	104.5	84.9	89.0	13.9
	24	103.7	94.0	86.3	94.7	8.7	107.8	131.1	93.5	110.8	19.0
DMC1	0	100.0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	100.0	0.0
	4	100.4	111.0	104.0	105.1	5.4	84.4	82.9	82.8	83.4	0.9
	8	125.2	132.9	118.0	125.4	7.5	79.7	87.3	86.9	84.6	4.3
	24	136.7	116.1	99.1	117.3	18.8	92.9	93.9	75.0	87.3	10.6

Table 3.7: The mRNA levels of *RAD51* and the 3 *RAD51*-like genes in response to 10 µg.ml-1 phleomycin. The results for the three repetitions of the DNA damage experiments (Expt 1, 2 and 3) carried out at 10 µg.ml-1 phleomycin are shown. The values represent the quantification of Northern blots probed with *RAD51*, *RAD51-3*, *RAD51-5* or *DMC1* (GENE) and their subsequent quantification after re-probing with *GPI-8* (GPI-8): the quantification value obtained for time 0 hours was regarded as 100% and all other values scaled accordingly. Averages (Avge) and the standard deviations (St dev) generated from the data are also shown.

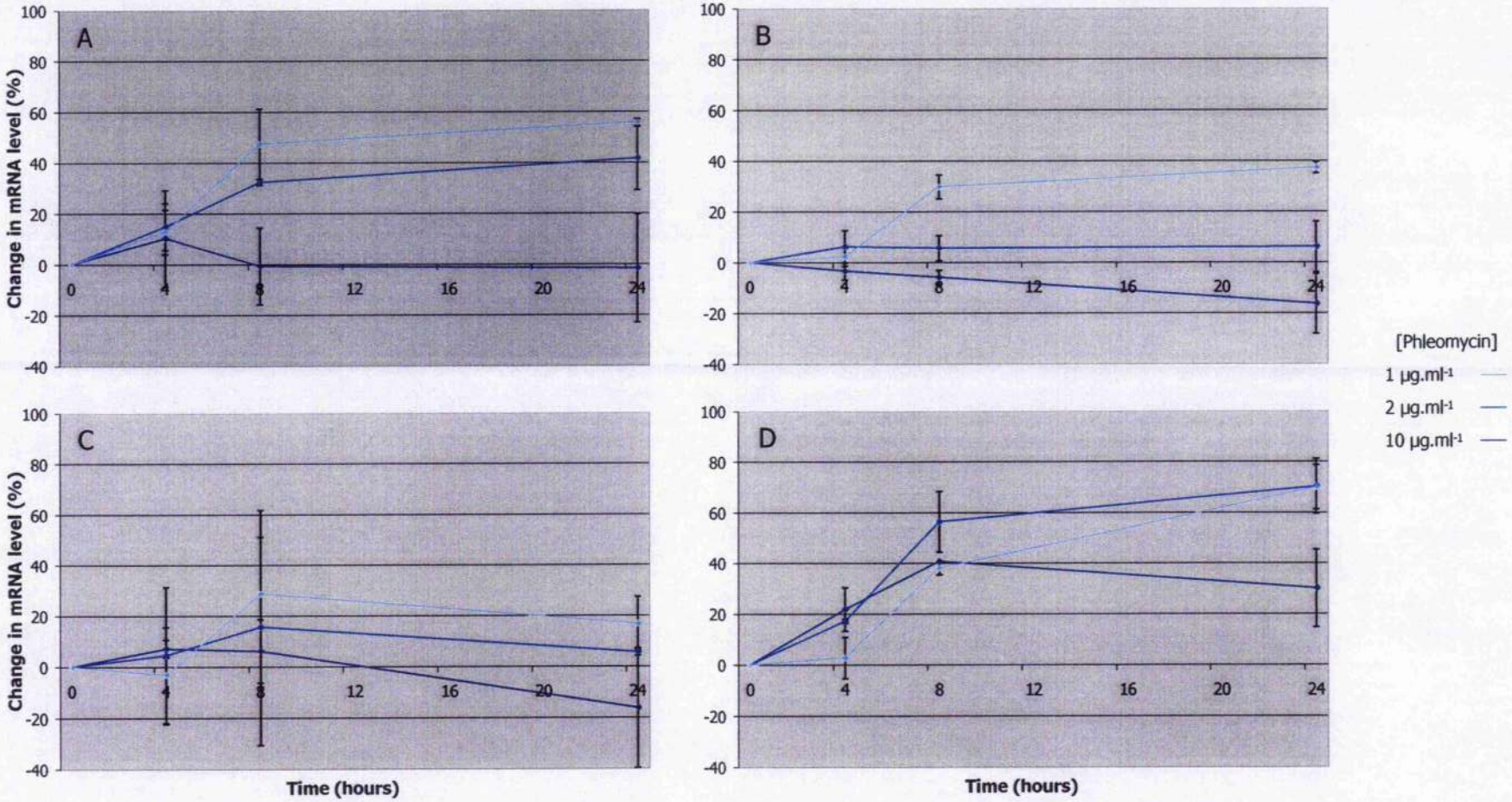


Figure 3.18: The change in *RAD51*, *RAD51-3*, *RAD51-5* and *DMC1* mRNA level in response to DNA damage. For each experiment the adjusted quantification values for *GPI-8* were subtracted from the adjusted quantification value for the gene of interest. This value represented the change in mRNA level in response to DNA damage relative to control mRNA. The average values obtained at each concentration of phleomycin [Phleomycin] are shown, as are the standard deviation. A: *RAD51*. B: *RAD51-3*. C: *RAD51-5*. D: *DMC1*.

No significant change could be detected for the *RAD51-5* mRNA levels in response to DNA damage. *RAD51-3* appeared to be similar, although at 1 $\mu g.ml^{-1}$ phleomycin a modest (~40%) increase may be apparent after 24 hours. The results for *RAD51* and *DMC1* appear to suggest that after exposure to 1 or 2 $\mu g.ml^{-1}$ of phleomycin their mRNA levels increased by approximately 40 -70% after 24 hours. This result was perhaps expected for *RAD51*, as it has been shown to be involved in DNA damage repair in *T. brucei* (McCulloch and Barry, 1999) but it should be noted that this is a very modest change. A putative increase in *DMC1* is surprising, in contrast, if this is functionally involved in meiosis-specific recombination. No evidence exists for DMC1 acting in DNA damage repair in mitotic cells. To examine this question further, the mRNA levels of the open reading frames upstream and downstream of *DMC1* were analysed (Fig 3.19). It is assumed that, as is the case for most of the *T. brucei* genome (Vanhamme and Pays, 1995), *DMC1* is part of a multi gene transcription unit encompassing the surrounding genes. The downstream open reading frame is predicted to encode a replication licensing factor of the MCM family (MCM; Forsburg, 2004), while the upstream open reading frame (P-ORF) has no identifiable homologues. Because the expression of a replication licensing factor may be up-regulated in response to DNA damage, the blots were also probed with the next open reading frame downstream, tubulin-tyrosine ligase (TT-LIG; Erck *et al.*, 2000), which was deemed unlikely to respond to DNA damage.

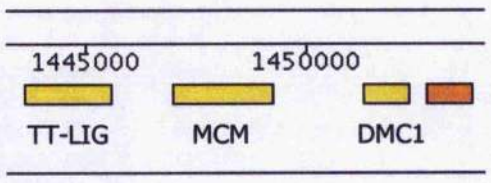


Figure 3.19: The open reading frames surrounding *DMC1*. The diagram, adapted from the *T. brucei* genome database, represents the genomic environment surrounding *DMC1*. Yellow boxes: putative open reading frame with predicted function. Orange box: putative open reading frame with no predicted function (referred to as P-ORF). TT-LIG: tubulin tyrosine ligase. MCM: replication licensing factor of the MCM family.

Regions of each of the open reading frames, approximately 600 bp in length, were PCR-amplified from Lister 427 genomic DNA using Taq DNA polymerase and the primers specific for each open reading frame (named 5′ and 3′ for each gene; appendix 1). The Northern blots of the experiments carried out at 10 $\mu g.ml^{-1}$ phleomycin were stripped and re-probed with PCR products and quantified as described previously. The results of this analysis are shown in Table 3.9 and graphed in Figure 3.20.

	probe	GENE					GPI-8				
	Time (hours)	Expt 1	Expt 2	Expt 3	Avge	St dev	Expt 1	Expt 2	Expt 3	Avge	St dev
MCM	0	100.0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	100.0	0.0
	4	122.7	148.7	120.1	130.5	15.8	84.4	82.9	82.8	83.4	0.9
	8	89.7	176.2	124.6	130.2	43.5	79.7	87.3	86.9	84.6	4.3
	24	210.4	230.4	117.7	186.1	60.1	92.9	93.9	75.0	87.3	10.6
TT LIG	0	100.0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	100.0	0.0
	4	164.4	183.4	150.1	166.0	16.7	131.3	93.2	87.8	104.1	23.7
	8	194.5	215.6	199.9	203.3	11.0	120.1	96.0	95.0	103.7	14.2
	24	162.2	291.0	191.4	214.9	67.5	117.4	122.6	99.2	113.1	12.3
P ORF	0	100.0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	100.0	0.0
	4	82.0	80.5	86.9	83.1	3.3	84.4	82.9	82.8	83.4	0.9
	8	87.7	79.1	95.5	87.4	8.2	79.7	87.3	86.9	84.6	4.3
	24	72.1	103.7	96.5	90.8	16.6	92.9	93.9	75.0	87.3	10.6

Table 3.8: The mRNA levels of the genes surrounding *DMC1* in response to 10 μg.ml^{-1} phleomycin. The results for the three repetitions of the DNA damage experiments (Expt 1, 2 and 3) carried out at 10 μg.ml^{-1} phleomycin are shown. The values represent the quantification of Northern blots probed with the DNA replication licensing factor (MCM), tubulin-tyrosine ligase (TT-LIG) and the putative open reading frame (P-ORF) and their subsequent quantification after re-probing with *GPI-8* (GPI-8). Quantification values obtained for time 0 hours was regarded as 100% and all other values scaled accordingly. Averages (Avge) and the standard deviations (St dev) generated from the data are also shown.

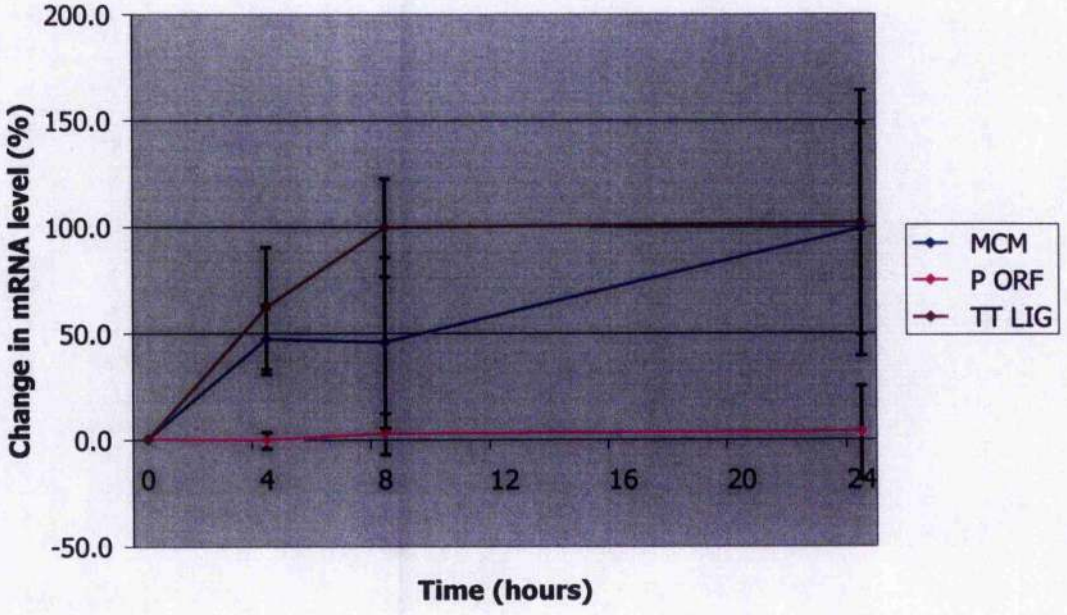


Figure 3.20: The change in the mRNA level of the genes surrounding *DMC1* response to 10 μg.ml^{-1} phleomycin. For each experiment the adjusted quantification values for *GPI-8* were subtracted from the adjusted quantification value for the gene of interest. This value represented the change in mRNA level in response to DNA damage relative to an unchanged control mRNA. The average values obtained for each open reading frame and the standard deviations are shown.

These results suggest that there was no change in the mRNA level for the putative open reading frame upstream of *DMC1* but that there was an approximate 100% increase (with significant variation) for both the putative DNA replication licensing factor and the putative tubulin tyrosine ligase after 24 hours. These values are greater than that observed for *RAD51* and *DMC1* and while it is possible to conceive that the mRNA level of a DNA replication factor would be increased in response to DNA damage, it is harder to conceive that the same would happen to a tubulin-tyrosinc ligase. The simplest explanation of them is that the modest increases in mRNA levels observed (approximately 40-100%) are not specific increases in response to phleomycin damage, but simply as a result of the variability of this assay. The increases, if they were real, were in any case relatively insignificant changes in mRNA level when compared to those observed in other organisms. It is clearly possible that DNA damage expression changes in *T. brucei* are orchestrated at translation or protein stability levels. It was therefore decided that the analysis of the mRNA level in response to DNA damage is not a suitable method to predict a role for a gene in DNA damage repair in *T. brucei*.

3.9 Summary

As a result of searching the *T. brucei* genome database and sequence comparisons, five further *RAD51*-like genes have been identified. One of them appears to encode the highly conserved enzyme Dmc1, whilst four others are more poorly conserved. All five genes have features consistent with RAD51-like function. So far, *T. brucei* is the only single-celled eukaryote outside of fungi to be examined for RAD51-related functions. The number of potential *RAD51*-like genes identified in *T. brucei*, however, is more akin to the number found in higher eukaryotes.

Three of the *RAD51*-like genes were chosen for further analysis, as analysis of all five was deemed beyond the scope of what is achievable during a PhD project. The remainder of this chapter concentrated on the initial characterisation of these genes. All three were shown to be single copy genes transcribed to putative 'mature' mRNA in Lister 427 bloodstream form cells. As antibodies were not available for any of the proteins, analysis of the mRNA level was analysed to assess the genes' potential to be involved in DNA damage repair. We found no evidence for an increase in mRNA level in response to DNA damage, suggesting that all three might not be involved in DNA damage repair. However, *RAD51*, a gene already shown to be involved in DNA damage

repair, also showed no clear increase in mRNA level in response to DNA damage. Therefore, it was concluded that analysis of the mRNA level in *T. brucei* is not a suitable method for the definition of a genes' function in DNA repair.

CHAPTER 4

ANALYSIS OF THE ROLES OF *RAD51-3*, *RAD51-5* AND *DMC1* IN DNA REPAIR, RECOMBINATION AND ANTIGENIC VARIATION.

4.1 Introduction

The only gene so far identified in *T. brucei* to have a role in antigenic variation is *RAD51*. *T. brucei RAD51* mutants have an increased sensitivity to DNA damaging agents, a recombination defect (with both reduced efficiency and aberrant integrations taking place), and a VSG switching defect with both gene conversion and *in situ* mechanisms affected. However, recombination and switching do still occur, and therefore back-up pathways must exist for both processes. In yeast (Paques and Haber, 1999) and humans (Wiese *et al.*, 2002), Rad51 operates with a number of cofactors or homologues that act to aid efficient recombination. As five *RAD51*-like genes have been discovered in *T. brucei*, the aim of this chapter is to analyse three of the *RAD51*-like genes to determine their role in DNA damage repair, recombination and VSG switching in bloodstream form cells.

4.2 Generation of gene disruption mutants in the 3174.2 strain of *T. brucei*

Homozygous mutants of the *RAD51*-like genes were generated using the same strategy that was used previously to generate *rad51* homozygous mutants in *T. brucei* (McCulloch and Barry, 1999). This approach does not create a classical 'knock out' mutant, where the whole of the open reading frame is removed, but a gene disruption where the core functional domains of the gene are removed. In this method 5′ and 3′ regions of the open reading frame were PCR-amplified, cloned and used as flanking sequence for recombination following transformation. The core of the gene was replaced with one of two antibiotic resistance markers, allowing the selection of constructs that have integrated into the genome and the disruption of both alleles (*T. brucei* is a diploid organism).

Mutation of *RAD51-3* and *RAD51-5* utilised this two-construct approach, but mutation of *DMC1* utilised a new method (see section 4.2.3). In all cases, the gene disruptions were designed to remove the central 5-8 core conserved domains, as discussed in chapter 3 (section 3.4). Moreover, all mutations were made in the *T. brucei* transgenic strain 3174.2 (McCulloch *et al.*, 1997; Rudenko *et al.*, 1996). This is a bloodstream stage derivative of Lister 427 (Melville *et al.*, 2000), which allows analysis of VSG switching frequency and mechanisms (section 4.8).

Oligonucleotide primers were designed for 5′ (primers LHF FOR and REV; Appendix 1) and 3′ (primers RHF-For and RHF-Rev; Appendix 1) regions of the *RAD51-3* and *RAD51-5* open reading frames (Fig. 4.1). The sequences were then PCR-amplified using Herculase DNA polymerase (Stratagene) and genomic DNA from 3174.2 *T. brucei* and subsequently cloned into pBC KS using the restriction sites contained within the primers (*Eco*RI, *Hind*III and *Xho*I; Fig. 4.1) resulting in the production of pCP301 (*RAD51-3* flanks) and pCP302 (*RAD51-5* flanks). Puromycin and blasticidin resistance cassettes, with flanking processing signals derived from the tubulin array to allow RNA trans-splicing (Vanhamme and Pays, 1995), were PCR-amplified using Herculase (primers βα5′ and αβ3′) and plasmids pTBT and pTPT as substrates (supplied by M. Cross; (Rudenko *et al.*, 1994)) and cloned into the *Hpa*I restriction site introduced between the 5′ and 3′ regions. This generated the constructs Δ*RAD51-3*::*PUR*, Δ*RAD51-5*::*PUR*, Δ*RAD51-3*::*BSD* and Δ*RAD51-5*::*BSD* (Figs. 4.2 and 4.3). The constructs were excised from pBC KS by restriction digestion with *Not*I and *Apa*I and the resulting digested DNA phenol:chloroform extracted and ethanol precipitated. Approximately 5 μg of digested DNA (quantified by gel electrophoresis relative to Life Technologies 1 kb size ladder) was then used in each *T. brucei* transformation.

4.2.1 Generation of *RAD51-3* mutants in the 3174.2 strain

Two transformations were carried out to generate two independent *RAD51-3* heterozygous cell lines using the Δ*RAD51-3*::*PUR* construct. To do this, 3174.2 cells were transformed using the protocol described in Section 2.4 and antibiotic-resistant transformants were selected for as described in Table 4.1. The generation of heterozygous mutants was confirmed by Southern analysis, carried out on *Eco*RI digested genomic DNA from puromycin resistant cell lines and probed with the *RAD51-3* LHF (data not shown; see sections 4.2.5 and 4.2.6). Subsequently, the two independent heterozygous mutants (*RAD51-3*+/- 1 and *RAD51-3*+/- 2) were transformed with the Δ*RAD51-3*::*BSD* construct in an attempt to generate two independent *rad51-3* homozygous mutants. Antibiotic transformants were selected as described in Table 4.1 and these transformations resulted in the generation of two putative *rad51-3* homozygous cells (*RAD51-3*-/- 1 and *RAD51-3*-/- 2). Selection for homozygous mutants using both antibiotics was unsuccessful, and therefore the second round transformants were selected initially using only selection for the transformed resistance gene (BSD).

Blasticidin-resistant transformants were subsequently placed on double drug selection before being considered a putative homozygous mutant (Table 4.1). The integration of the BSD marker was screened by PCR-amplification of the complete *RAD51-3* open reading frame, using primers *RAD51-3*-For and *RAD51-3*-Rev (Appendix 1) and Taq DNA polymerase, in double resistant cell lines (data not shown; see sections 4.2.5 and 4.2.6). Despite a large number of double resistant transformant cell lines being generated (Table 4.1), only a small number of them were found to be putative *rad51-3-/-* mutants. This is probably as a result of integration of the blasticidin resistance gene into the tubulin array using homology found within Δ*RAD51-3*::*BSD*.

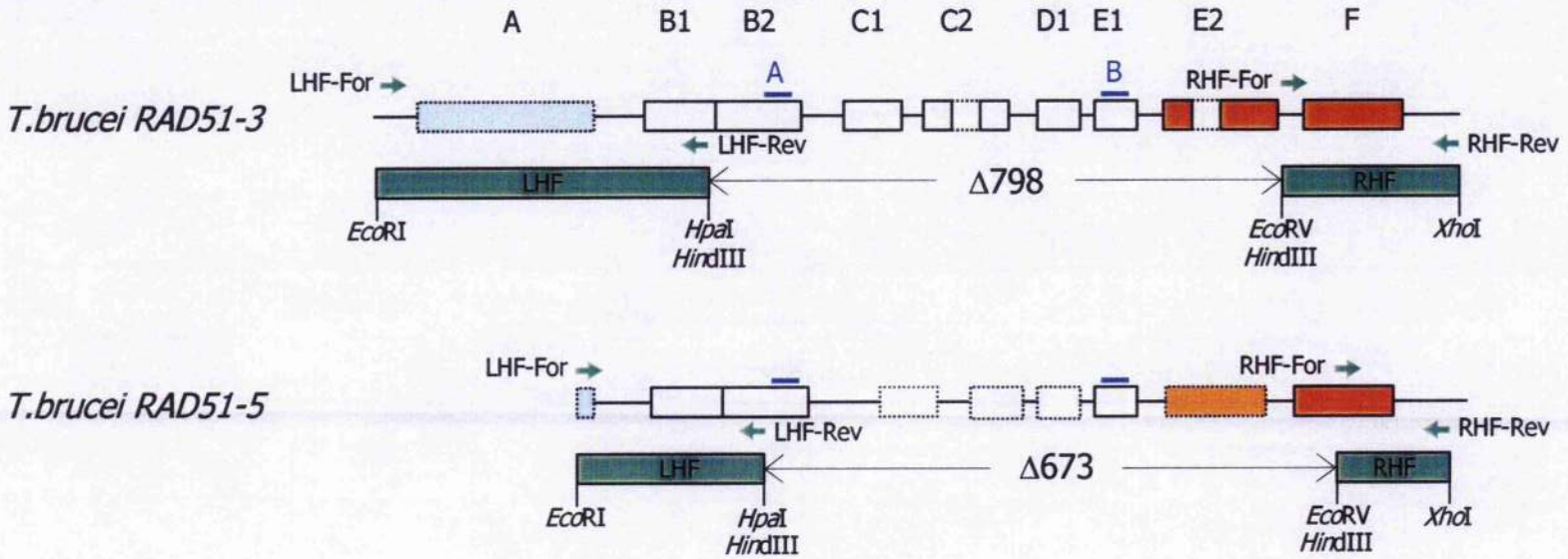


Figure 4.1: The 5′ and 3′ regions of the *RAD51-3* and *RAD51-5* open reading frames used in the gene disruption constructs. The *RAD51-3* primers LHF-For and LHF-Rev amplify a 350 base pair 5′ region (RHF) of the open reading frame and primers RHF-For and RHF-Rev amplify a 370 base pair 3′ region (RHF). Using these regions in the disruption constructs for the targeting of the resistance cassettes, which replaces the sequence between the 5′ and 3′ regions, results in the removal of 798 bp containing 6 of the core domains of *RAD51-3*, including the highly conserved walker A (A) and B (B) boxes. The *RAD51-5* primers LHF-For and LHF-Rev amplify a 200 base pair 5′ region (RHF) of the open reading frame and primers RHF-For and RHF-Rev amplify a 230 base pair 3′ region (LHF). Using these regions in the disruption constructs for the targeting of the resistance cassettes, which replaces the sequence between the 5′ and 3′ regions, results in the removal of 673 bp containing 5 of the core domains of *RAD51-5*, including the highly conserved walker A (A) and B (B) boxes. In both cases the primers contain restriction sites (listed) to aid cloning into pBC KS and the subsequent cloning of the resistance cassette between the flanks. LHF: left hand flank (homologous to the 5′ region of the open reading frame). RHF: right hand flank (homologous to the 3′ region of the open reading frame).

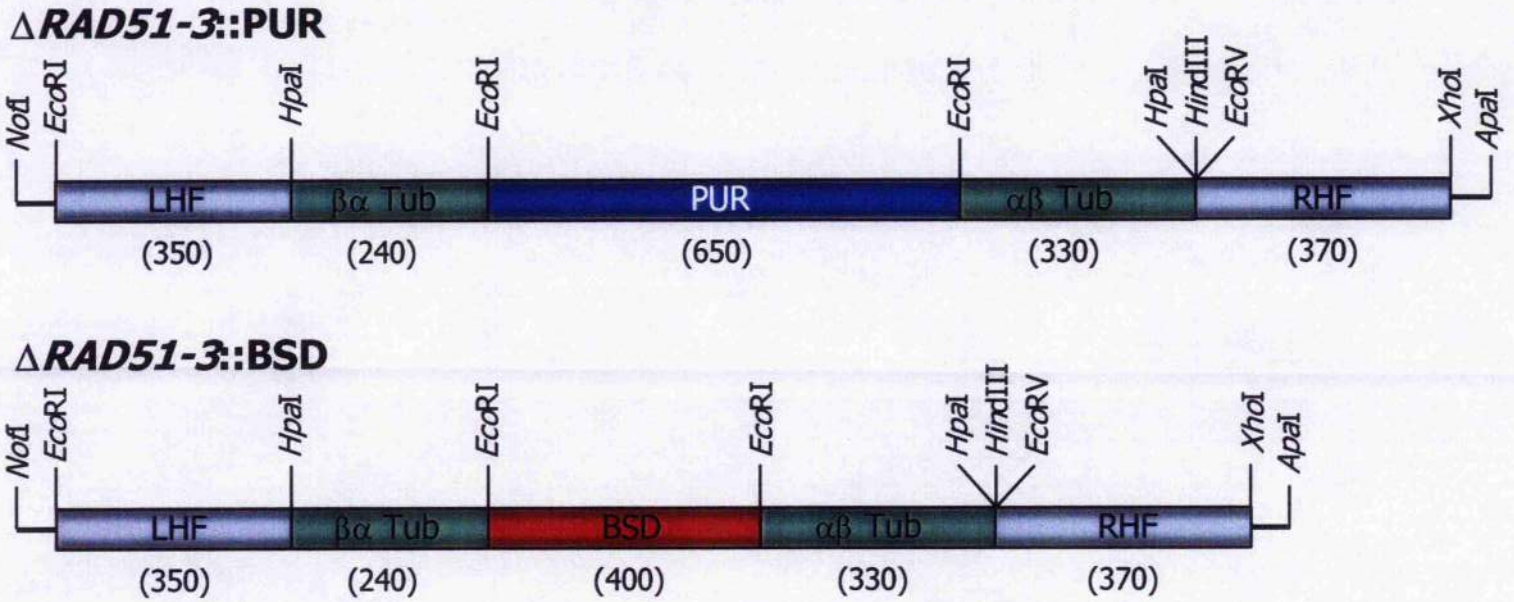


Figure 4.2: *RAD51-3* gene disruption constructs. Restriction maps of the constructs used for the disruption of *RAD51-3*. Sizes of the individual components are shown in base pairs. Constructs were cloned into the pBC KS plasmid. LHF: left hand flank (homologous to the 5' region of the open reading frame). βα Tub: βα Tubulin intergenic region (processing signal). PUR: puromycin resistance gene open reading frame. αβ Tub: αβ Tubulin intergenic region (processing signal). RHF: right hand flank (homologous to the 3' region of the open reading frame). BSD: blasticidin resistance gene open reading frame.

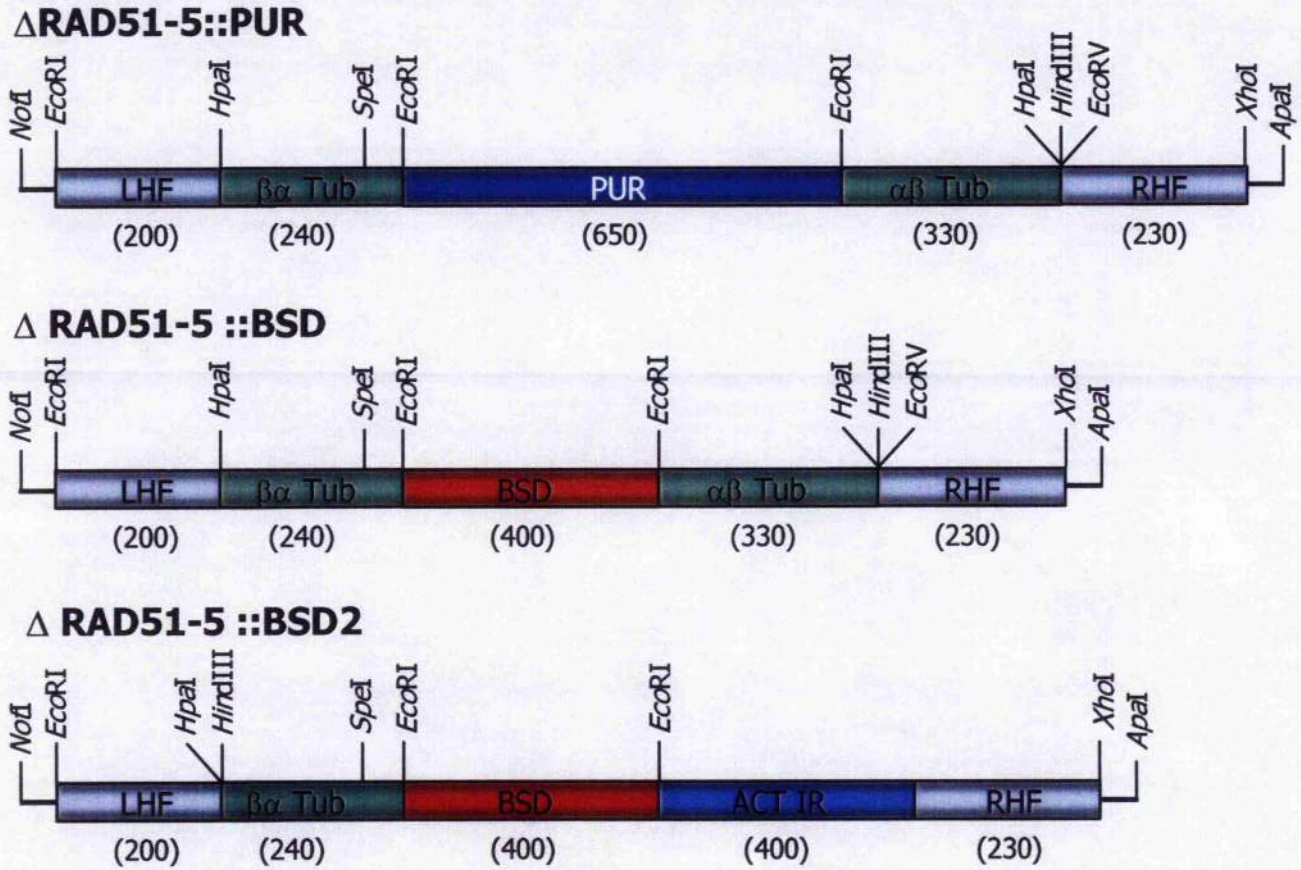


Figure 4.3: *RAD51-5* gene disruption constructs. Restriction maps of the constructs used for the disruption of *RAD51-5*. Sizes of the individual components are shown in base pairs. Constructs were cloned into the pBC KS plasmid. ACT IR: Actin intergenic region (processing signal); all other components are described in Figure 4.2.

Cell line being generated	Trans. Number	Construct used	Cell No. plated out	Drug(s) and concentration(s)	No. of transformants	No of correct integrations
Heterozygote	1	Δ*RAD51-3*::*PUR*	1.0 X 10^7	PUR 1.0	4	3
		Δ*RAD51-3*::*PUR*	1.0 X 10^7	PUR 0.5	3	X
	2	Δ*RAD51-3*::*PUR*	1.0 X 10^7	PUR 1.0	2	2
		Δ*RAD51-3*::*PUR*	1.0 X 10^7	PUR 0.5	4	X
Homozygote 1	1	Δ*RAD51-3*::*BSD*	2.0 X 10^7	BSD 2.5	5	0
	2	Δ*RAD51-3*::*BSD*	2.0 X 10^7	BSD 1.75 & PUR 0.5	0	
Homozygote 2	1	Δ*RAD51-3*::*BSD*	2.0 X 10^7	BSD 2.5	3	2
	2	Δ*RAD51-3*::*BSD*	2.0 X 10^7	BSD 1.75 & PUR 0.5	5	0
Homozygote 1	5	Δ*RAD51-3*::*BSD*	1.0 X 10^7	BSD 2.5	4	0
	6	Δ*RAD51-3*::*BSD*	1.0 X 10^7	BSD 2.5	0	
Homozygote 1	7	Δ*RAD51-3*::*BSD*	3.2 X 10^7	BSD 2.5	38	0
	8	Δ*RAD51-3*::*BSD*	3.3 X 10^7	BSD 2.5	39	0
Homozygote 1	9	Δ*RAD51-3*::*BSD*	1.9 X 10^7	BSD 2.5	24 (8)	3
	10	Δ*RAD51-3*::*BSD*	1.6 X 10^7	BSD 2.5	18 (8)	3
	11	Δ*RAD51-3*::*BSD*	1.4 X 10^7	BSD 2.5	18 (8)	1

Table 4.1: Transformations carried out during the generation of *RAD51-3* mutants. A list of the transformations carried out during the generation of two independent *RAD51-3* heterozygous and homozygous mutants. The table includes the constructs and drug selection used, the numbers of cells plated out, the number of those that were antibiotic resistant transformants and the number that had correctly integrated the construct. The generation of homozygous mutant cell lines selection initially utilised only selection for the transformed resistance gene - blasticidin (BSD), with the exception of homozygote transformations 2. Transformants were subsequently placed on double drug selection (blasticidin as listed and puromycin at 0.5 $\mu g.ml^{-1}$) before being considered a putative homozygous mutant. All drug concentrations listed are in $\mu g.ml^{-1}$.Trans.: transformation. No.: Number. PUR: puromycin. X: transformants not analysed. (8): 8 of the transformants were analysed.

4.2.2 Generation of *RAD51-5* mutants in the 3174.2 strain

Two transformations were carried out to generate two independent *RAD51-5* heterozygous cell lines using the Δ*RAD51-5*::*PUR* construct. Again, 3174.2 cells were transformed (see protocol section 2.4), antibiotic resistant transformants selected for as described in table 4.2 and screened by PCR-amplification of the entire *RAD51-5* open reading frame. After 4 attempts, no putative *RAD51-5* heterozygote mutants were obtained, with all transformants having integrated the PUR cassette into locations other than the *RAD51-5* open reading frame. The drug concentration was, therefore, reduced and 4 subsequent transformations resulted in the generation of two independent putative heterozygous mutants. The generation of putative heterozygous mutants (*RAD51-5* +/- 1 and *RAD51-5*+/- 2) was confirmed by PCR-amplification of the entire *RAD51-5* open reading frame using primers *RAD51-5*-For and *RAD51-5*-Rev (Appendix 1) and Taq DNA polymerase (data not shown; see sections 4.2.5 and 4.2.6). Subsequent transformation of the putative *RAD51-5*+/- 2 using the Δ*RAD51-5*::BSD construct resulted in the generation of a putative *rad51-5* homozygous mutant (Table 4.2), confirmed by PCR as described above, although most transformants again had integrated the BSD cassette outside of the *RAD51-5* open reading frame. In contrast, 8 transformations of the putative *RAD51-5*+/- 1 cell line failed to generate an independent *rad51-5* homozygous mutant (Table 4.2). The blasticidin resistance cassette was always integrated into locations other than the *RAD51-5* open reading frame, producing cells that were double resistant to both puromycin and blasticidin but were still heterozygous (data not shown). To overcome this problem, the Δ*RAD51-5*::*BSD* construct was redesigned to create Δ*RAD51-5*::*BSD2* (Fig 4.3). In this construct, the blasticidin resistance gene was flanked by an actin intergenic region as the 3′ processing signal, thereby preventing integration into tubulin.

Cell line being generated	Trans. number	Construct used	Cell No. plated out	Drug and concentration	No. of transformants	No of correct integrations
Heterozygote	1	Δ*RAD51-5*::*PUR*	1.0 X 10^7	PUR 1.0	2	0
			1.0 X 10^7	PUR 0.5	2	0
Heterozygote	2	Δ*RAD51-5*::*PUR*	1.0 X 10^7	PUR 1.0	4	0
			1.0 X 10^7	PUR 0.5	2	0
Heterozygote	3	Δ*RAD51-5*::*PUR*	1.0 X 10^7	PUR 0.5	7	0
			1.0 X 10^7	PUR 0.25	4	0
Heterozygote	4	Δ*RAD51-5*::*PUR*	1.0 X 10^7	PUR 0.5	0	
			1.0 X 10^7	PUR 0.25	0	
Heterozygote	5	Δ*RAD51-5*::*PUR*	1.0 X 10^7	PUR 0.5	0	
			1.0 X 10^7	PUR 0.25	2	1 C
Heterozygote	6	Δ*RAD51-5*::*PUR*	1.0 X 10^7	PUR 0.5	4	1
			1.0 X 10^7	PUR 0.25	1	1
Heterozygote	7	Δ*RAD51-5*::*PUR*	2.7 X 10^7	PUR 0.25	14	0
	8	Δ*RAD51-5*::*PUR*	3.3 X 10^7	PUR 0.25	12	3
Homozygote 2	1	Δ*RAD51-5*::*BSD*	1.0 X 10^7	BSD 2.5	12	1
	2	Δ*RAD51-5*::*BSD*	1.0 X 10^7	BSD 2.5	12	0
Homozygote 1	1	Δ*RAD51-5*::*BSD*	2.3 X 10^7	BSD 2.5	27	0
	2	Δ*RAD51-5*::*BSD*	2.9 X 10^7	BSD 2.5	35	0
	3	Δ*RAD51-5*::*BSD*	2.3 X 10^7	BSD 2.5	27	0
Homozygote 1	4	Δ*RAD51-5*::*BSD*	1.3 X 10^7	BSD 2.5	18	0
	5	Δ*RAD51-5*::*BSD*	1.7 X 10^7	BSD 2.5	21	0
	6	Δ*RAD51-5*::*BSD*	1.5 X 10^7	BSD 2.5	18	0
Homozygote 1	7	Δ*RAD51-5*::*BSD2*	2.0 X 10^7	BSD 2.5	6	0
	8	Δ*RAD51-5*::*BSD2*	2.0 X 10^7	BSD 2.5	12	0
Homozygote 2	3	Δ*RAD51-5*::*BSD2*	2.0 X 10^7	BSD 2.5	11	0
	4	Δ*RAD51-5*::*BSD2*	2.0 X 10^7	BSD 2.5	8	0
Homozygote 1	9	Δ*RAD51-5*::*BSD2*	3.6 X 10^7	BSD 2.5	10	0
	10	Δ*RAD51-5*::*BSD2*	3.0 X 10^7	BSD 2.5	15	0
Homozygote 2	5	Δ*RAD51-5*::*BSD2*	4.0 X 10^7	BSD 2.5	16	0
	6	Δ*RAD51-5*::*BSD2*	3.6 X 10^7	BSD 2.5	17	1
Heterozygote	9	Δ*RAD51-5*::*PUR*	4.0 X 10^7	PUR 0.25	4	0
	10	Δ*RAD51-5*::*PUR*	3.9 X 10^7	PUR 0.25	4	1
Homozygote 3	1	Δ*RAD51-5*::*BSD2*	3.6 X 10^7	BSD 2.5	12	2
	2	Δ*RAD51-5*::*BSD2*	2.6 X 10^7	BSD 2.5	7	X
	3	Δ*RAD51-5*::*BSD2*	3.6 X 10^7	BSD 2.5	15	X

Table 4.2: Transformations carried out during the generation of *RAD51-5* mutants. A list of the transformations carried out during the generation of two independent *RAD51-5* heterozygous and homozygous mutants. The table includes the constructs and drug selection used, the numbers of cells plated out, the number of those that were antibiotic resistant transformants and the number that had correctly integrated the construct. During the generation of homozygous mutant cell lines selection initially utilised only selection for the transformed resistance gene - blasticidin (BSD). Transformants were subsequently placed on double drug selection (blasticidin as listed and puromycin at 0.25 $\mu g.ml^{-1}$) before being considered a putative homozygous mutant. All drug concentrations listed are in $\mu g.ml^{-1}$. Trans.: transformation. No.: Number. PUR: puromycin. C: transformant lost due to contamination. X: transformants not analysed.

To generate Δ*RAD51-5*::*BSD2*, the new blasticidin resistance cassette, with flanking tubulin and actin processing signals, was PCR-amplified from the construct pCP101 (Conway *et al.*, 2002b) using Herculase and the primers βα5′-*Hpa*I and ACT3′-*Sph*I (Appendix 1). Due to the lack of appropriate restriction sites available for cloning the blasticidin cassette into pCP302, the ends of the PCR product were made blunt using the Klenow fragment of *E. coli* DNA polymerase (New England Biolabs) and ligated into pCP302, which had been digested with *Eco*RV and treated with CIP (New England Biolabs), to create Δ*RAD51-5*::*BSD2* (Fig 4.3). DNA of this construct was prepared for transformation as described in section 4.2.

It was decided that both putative heterozygous mutants should be transformed using the new second round construct, despite the fact that one of them had already been used to generate a homozygous mutant, to ensure that both independent homozygous cell lines were equivalent. Transformations of the two putative *RAD51-5*+/- cell lines using Δ*RAD51-5*::*BSD2* were carried out (as per protocol, Section 2.4) and antibiotic resistant transformants selected for as described in Table 4.2. Four transformations were carried out on each cell line which resulted in the generation of only one putative *rad51-5* homozygous mutant (*RAD51-5*-/- 2), confirmed by PCR as described above (data not shown; see Sections 4.2.5 and 4.2.6). Further analysis of the putative *RAD51-5* heterozygous cell line, *RAD51-5*+/- 1, showed that its genome contained two puromycin resistance markers, one in the disrupted *RAD51-5* allele and another in the tubulin array (data not shown). Transformation of this cell line using the Δ*RAD51-5*::*BSD2* construct resulted in it replacing the original disruption construct resulting in a cell line that was double resistant to both blasticidin and puromycin but only *RAD51-5*+/-. As a result, the generation of an independent *rad51-5* homozygous cell line had to involve the generation of a new heterozygous cell line. The generation of a *RAD51-5* heterozygous cell line, using Δ*RAD51-5*::*PUR*, and a *rad51-5* homozygous cell line, using Δ*RAD51-5*::*BSD2*, was carried out as described previously in this section. The transformations (table 4.2) resulted in the generation of a putative *RAD51-5* heterozygous cell line (*RAD51-5* +/- 1) and a putative *rad51-5* homozygous cell line (*RAD51-5* -/- 1), confirmed by PCR as described above (data not shown; see sections 4.2.5 and 4.2.6).

4.2.3 Generation of *DMC1* mutants in the 3174.2 strain

A new method of gene disruption was attempted for the generation of *DMC1* mutants. This approach is based upon the work of Gaud *et al.* (1997) and Shen *et al.* (2001), who showed that homology-based integration of DNA can be achieved in *T. brucei* using as little as 42 bp of sequence homology. Testing this method of gene disruption would allow us to assess its suitability for routine use in the laboratory, as it would greatly reduce the time taken to generate mutants in the future. We therefore used PCR to amplify the blasticidin and bleomycin resistant cassettes and incorporated 50 bases of sequence derived from the 5′ and 3′ end of the DMC1 open reading frame (Fig 4.4) into the two PCR primers (*DMC1*-PCR-KO5′ and *DMC1*-PCR-KO3′; Appendix 1). PCR amplification using Herculase DNA polymerase and the plasmids pCP101 (Conway *et al.*, 2002b) and pRM481 (R. McCulloch, gift) as templates generated the products diagrammed in Figure 4.5, named Δ*DMC1::BSD* and Δ*DMC1::BLE*. Multiple PCRs were performed, the products pooled, purified using the Strataprep PCR purification kit (Stratagene) and approximately 5 μg of purified PCR product (quantified by gel electrophoresis relative to Life Technologies 1 kb size ladder) used in each transformation. Two transformations were carried out to generate two independent *DMC1* heterozygous cell lines using the Δ*DMC1::BSD* PCR product. Again, 3174.2 cells were transformed (see protocol section 2.4), antibiotic resistant transformants selected for as described in Table 4.3 and screened by PCR. After two attempts no heterozygous mutants were generated, with integration of the construct occurring into unknown locations. Six subsequent transformations using a reduced drug concentration for selection resulted in the generation of putative heterozygous mutants (Table 4.3). Screens for the generation of heterozygous mutants was carried out by PCR-amplification of the entire *DMC1* open reading frame using primers *DMC1*-For and *DMC1*-Rev (Appendix 1) and Taq DNA polymerase (data not shown; see Sections 4.2.5 and 4.2.6).

Subsequent transformation of the putative *DMC1* heterozygous cell lines (named *DMC1*+/- 1 and *DMC1*+/- 2) using the Δ*DMC1::BLE* PCR product, however, failed to result in the generation of *dmc1* homozygous mutants (Table 4.3). In 10 transformations, 266 transformants were screened by PCR (as described above) and none appeared to have integrated into the *DMC1* locus and have presumably recombined into non-homologous or partly homologous loci.

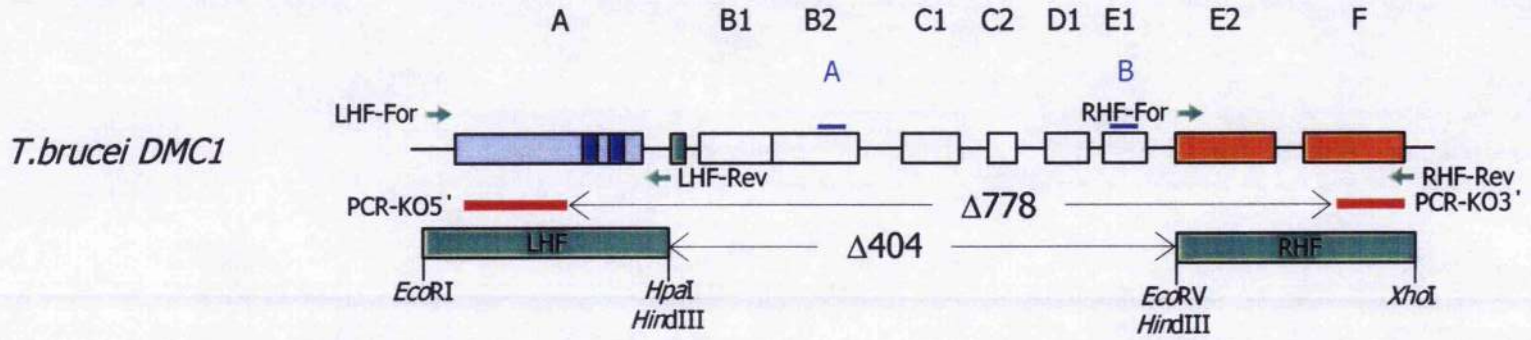


Figure 4.4: The 5′ and 3′ regions of the *DMC1* open reading frame used in the gene disruption constructs. The *DMC1* primers PCR-KO5′ and PCR-KO3′ contain 50 bases of homology (red line) to 5′ and 3′ regions of the open reading frame. Using these regions in the disruption constructs for the targeting of the resistance cassette, which replaces the sequence between the 5′ and 3′ regions, results in the removal of778 bp containing 8 of the core domains of *DMC1*, including the highly conserved walker A (A) and B (B) boxes. The *DMC1* primers LHF-For and LHF-Rev amplify a 260 base pair 5′ region of region of the open reading frame and primers RHF-For and RHF-Rev amplify a 250 base pair 3′ region. Using these regions in the disruption constructs for the targeting of the resistance cassette, which replaces the sequence between the 5′ and 3′ regions, results in the removal of 404 bp containing 7 of the core domains of *DMC1*, including the highly conserved walker A (A) and B (B) boxes. The primers contain restriction sites (listed) to aid cloning into pBC KS and the subsequent cloning of the resistance cassette between the flanks. LHF: left hand flank (homologous to the 5′ region of the open reading frame). RHF: right hand flank (homologous to the 3′ region of the open reading frame).

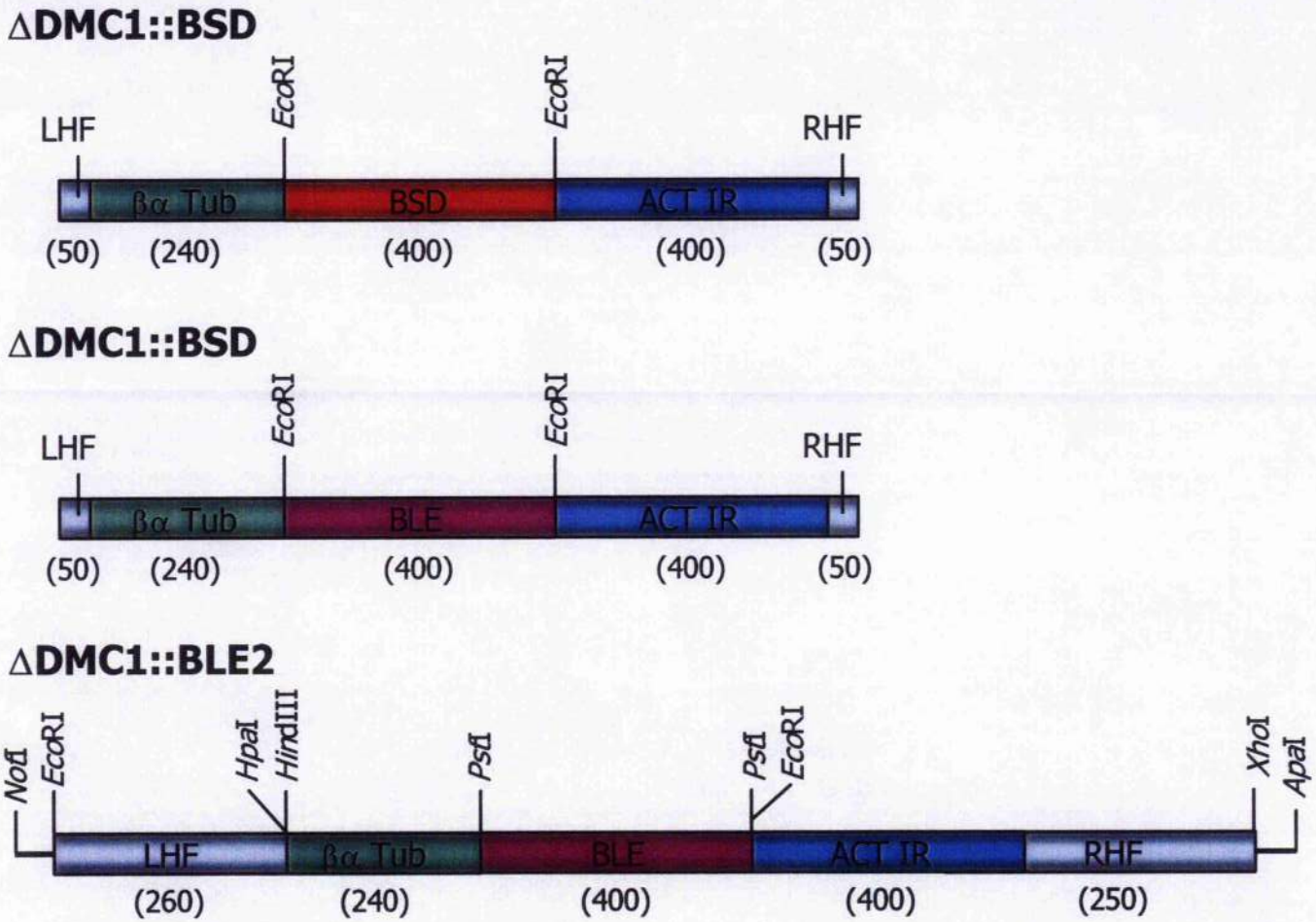


Figure 4.5: *DMC1* gene disruption constructs. Restriction maps of the constructs used for the disruption of the *DMC1* open reading frame. Sizes of the individual components are shown in base pairs. Note that the Δ*DMC1*::*BSD* and Δ*DMC1*::*BLE* are PCR products, while the Δ*DMC1*::*BLE2* construct was cloned into the pBC KS plasmid. BLE: Bleomycin resistance cassette. All other components labelled as in Figures 4.2 and 4.3.

Cell line being generated	Trans. Number	Construct used	Cell No. plated out	Drug and concentration	No. of transformants	No. of correct integrations
Heterozygote	1	Δ*DMC1*::*BSD*	1.0 X 10^7	BSD 10	0	
			1.0 X 10^7	BSD 5	3	0
	2	Δ*DMC1*::*BSD*	1.0 X 10^7	BSD 10	3	0
			1.0 X 10^7	BSD 5	0	
Heterozygote	3	Δ*DMC1*::*BSD*	1.8 X 10^7	BSD 10	5	0
			1.8 X 10^7	BSD 2.5	5	0
	4	Δ*DMC1*::*BSD*	1.7 X 10^7	BSD 10	3	0
			1.7 X 10^7	BSD 2.5	4	4
Heterozygote	5	Δ*DMC1*::*BSD*	2.0 X 10^7	BSD 2.5	0	
	6	Δ*DMC1*::*BSD*	2.0 X 10^7	BSD 2.5	0	
Heterozygote	7	Δ*DMC1*::*BSD*	2.0 X 10^7	BSD 2.5	0	
Heterozygote	8	Δ*DMC1*::*BSD*	2.9 X 10^7	BSD 2.5	11	0
	9	Δ*DMC1*::*BSD*	2.2 X 10^7	BSD 2.5	6	1
Homozygote 2	1	Δ*DMC1*::*BLE*	2.4 X 10^7	BLE 0.5	17	0
	2	Δ*DMC1*::*BLE*	2.0 X 10^7	BLE 0.5	15	0
Homozygote 1	1	Δ*DMC1*::*BLE*	1.0 X 10^7	BLE 0.5	0	
	2	Δ*DMC1*::*BLE*	1.0 X 10^7	BLE 0.5	0	
Homozygote 1	3	Δ*DMC1*::*BLE*	3.0 X 10^7	BLE 1.0	36	0
	4	Δ*DMC1*::*BLE*	2.9 X 10^7	BLE 1.0	34	0
Homozygote 2	3	Δ*DMC1*::*BLE*	3.2 X 10^7	BLE 1.0	27	0
	4	Δ*DMC1*::*BLE*	2.7 X 10^7	BLE 1.0	25	0
	5	Δ*DMC1*::*BLE*	3.2 X 10^7	BLE 1.0	30	0
	6	Δ*DMC1*::*BLE*	2.4 X 10^7	BLE 1.0	26	0
Homozygote 1	5	Δ*DMC1*::*BLE2*	3.0 X 10^7	BLE 1.0	20	6
Homozygote 2	6	Δ*DMC1*::*BLE2*	2.2 X 10^7	BLE 1.0	13	2

Table 4.3: Transformations carried out during the generation of *DMC1* mutants. A list of the transformations carried out during the generation of two independent *DMC1* heterozygous and homozygous mutants. The table includes the constructs and drug selection used, the numbers of cells plated out, the number of those that were antibiotic resistant transformants and the number that had correctly integrated the construct. The generation of homozygous mutant cell lines selection initialy utilised only selection for the transformed resistance gene - bleomycin (BLE). Transformants were subsequently placed on double drug selection (phleomycin (PLE) as listed and blasticidin at 2.5 $\mu g.ml^{-1}$) before being considered a putative homozygous mutant. All drug concentrations listed are in $\mu g.ml^{-1}$. Trans.: transformation. No.: Number. BSD: Blasticidin.

Because of this, a gene disruption construct was made for mutating the second *DMC1* allele using the traditional method. The re-designed construct is essentially the same as the PCR based product, with the exception that it contains larger targeting flanks of *DMC1* sequence to aid integration of the construct into the correct location.

Primers were designed to amplify 5′ (primers *DMC1*-LHF-For and *DMC1*-LHF-Rev; Appendix 1) and 3′ (primers *DMC1*-RHF-For and *DMC1*-RHF-Rev; Appendix 1) regions of the *DMC1* open reading frame. The 5′ and 3′ regions, 260 and 250 base pairs in size respectively, were PCR-amplified with Herculase DNA polymerase from 3174.2 genomic DNA and subsequently cloned into pBC KS using the restriction sites contained within the primers (*Eco*RI, *Hind*III and *Xho*I; Fig 4.4) resulting in the generation of pCP303. The bleomycin resistance cassette, with flanking tubulin and actin processing signals, was PCR-amplified from the construct pRM481 (R. McCulloch, gift) using Herculase DNA polymerase and the primers βα5′-*Hpa*I and ACT3′-*Sph*I. Due to the lack of appropriate restriction sites available for cloning the bleomycin resistance cassette between the targeting flanking sequences, the ends of the PCR product were made blunt using the Klenow fragment of *E. coli* DNA polymerase I and ligated into pCP303 (which had been digested with *Eco*RV and treated with CIP) to create Δ*DMC1*::*BLE2* (Fig 4.5). DNA of this construct was prepared for transformation as described in section 4.2.

Transformations of the putative *DMC1*+/- 1 and *DMC1*+/- 2 cell lines using Δ*DMC1*::*BLE2* were carried out as described in Section 2.4 and antibiotic resistant transformants selected for as described in Table 4.3. One transformation was carried out on each cell line which resulted in the generation of 33 double resistant cell lines which were screened by PCR (as described above) showing 8 to be putative homozygous mutants (data not shown; see Sections 4.2.5 and 4.2.6), two of which were chosen (*DMC1*-/- 1 and *DMC1*-/- 2), one from each *DMC1*+/- cell line.

4.2.4 Sub-cloning of the *RAD51-3*, *RAD51-5* and *DMC1* mutants

Due to the time in culture differences between each of the heterozygous mutants, homozygous mutants and the wild type cells, each cell line was sub-cloned to ensure that, for each gene, all cell lines during analysis had been in culture for the same time post sub-cloning. For all three genes, each of the heterozygous and homozygous mutants along with wild type cells, previously grown in non-selective media to 10^6

cells.ml^{-1}, were plated out in selective media at 0.5 cells per well in 96 well plates. All of the cell lines were plated out in media containing hygromycin (5.0 μg.ml^{-1}) and G418 (2.5 μg.ml^{-1}), and in addition, each of the heterozygous and homozygous mutants cell lines were plated out in media also containing the same antibiotics, at the same concentration, used during their generation. For each cell line three of the first wells to show growth were selected, grown up and stabilated. Stabilates from one of these selected cell lines were then used during further analysis of the *RAD51*-like genes, using a fresh stabilate for each assay.

4.2.5 Confirmation of the generation of mutants by Southern analysis

To confirm the generation of two independent heterozygous and homozygous mutants for each gene, Southern analysis was carried out on genomic DNA from each cell line, along with the 3174.2 parent. Genomic DNA was prepared from 25 ml of each cell line grown to approximately 2 x 10^6 cell.ml^{-1}. 5μg of DNA from each cell line was restriction digested for 6 hrs before being Southern blotted. The blots were then probed with the entire open reading frame of the appropriate gene, amplified by PCR from 3174.2 genomic DNA using Taq DNA polymerase and primers Exp-For and Exp-Rev (see Appendix 1). The restriction enzymes used, and the expected fragment sizes, for each cell line are in listed in Table 4.4 (sizes are shown in base pairs). Fragment sizes were calculated from the gene sequences (Appendices 3-5), the gene disruption constructs (Figs. 4.2-3 and 4.5), the genome sequence surrounding the genes (Fig. 4.6) and is summarised in Figures 4.7-9.

Gene	Enzyme	Cell line		
		Wild type	+/-	-/-
RAD51-3	*Bst*XI	2404	3109 2404	3109 2857
RAD51-5	*Spe*I	5465	5465 4077 1940	4077 1940 1760
DMC1	*Eco*RI	3010	3010 1991 860	2180 1991 1475 860

Table 4.4: Predicted fragment sizes for Southern analysis. These are the predicted fragment sizes expected when genomic DNA from heterozygous and homozygous mutants for each gene and wild-type cells is digested with the appropriate enzyme (listed). Fragment sizes, shown in base pairs, were calculated from the gene sequences (Appendices 3-5), the gene disruption constructs (Figs 4.2-3 and 4.5), the genomic environment surrounding each gene (Fig 4.6) predicted by the *T. brucei* genome sequences (gene DB at Sanger) and is summarised in Figures 4.7-9.

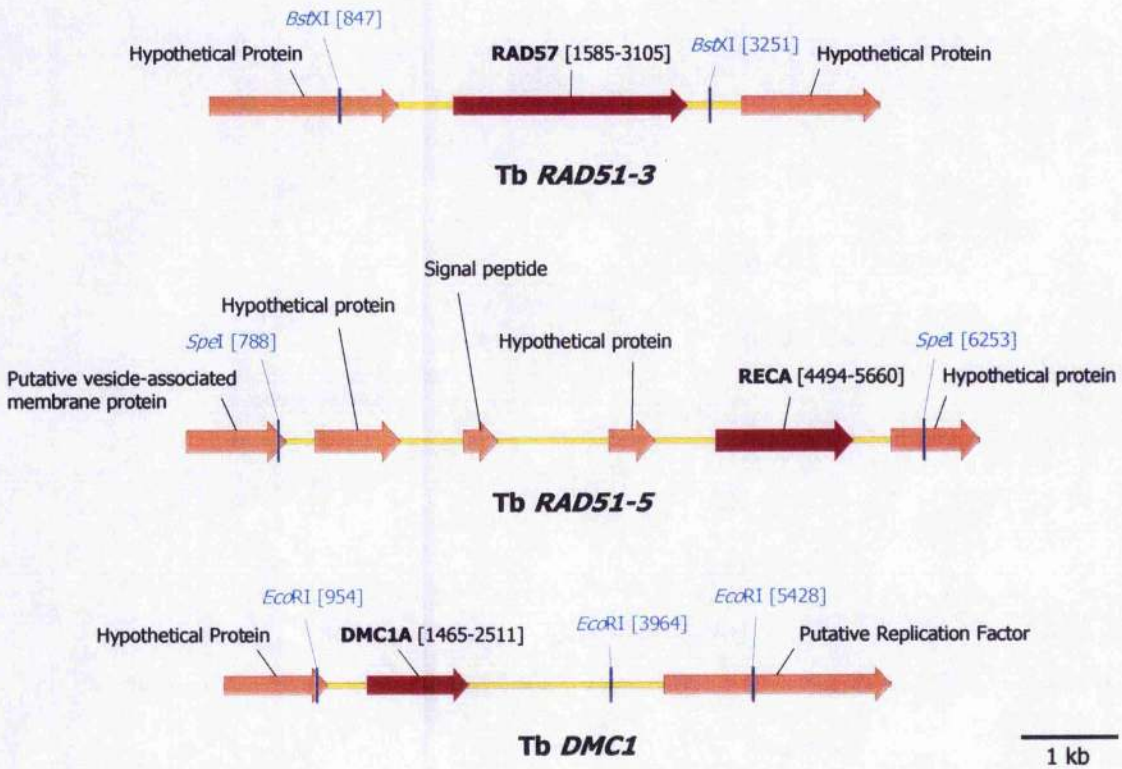


Figure 4.6: The genomic environment surrounding *RAD51-3*, *RAD51-5* and *DMC1*. Diagrams representing the genomic environment surrounding the *RAD51-3*, *RAD51-5* and *DMC1* open reading frames. The genome sequences were derived from the *T. brucei* genome database (compiled by Sanger and TIGR) and visualised using Vector NTI. The diagrams include the restriction enzyme sites, and the ORF positions, shown in bp relative to the sequence shown (figures in square brackets) used to calculate the predicted fragment sizes for Southern analysis of the heterozygous and homozygous mutants.

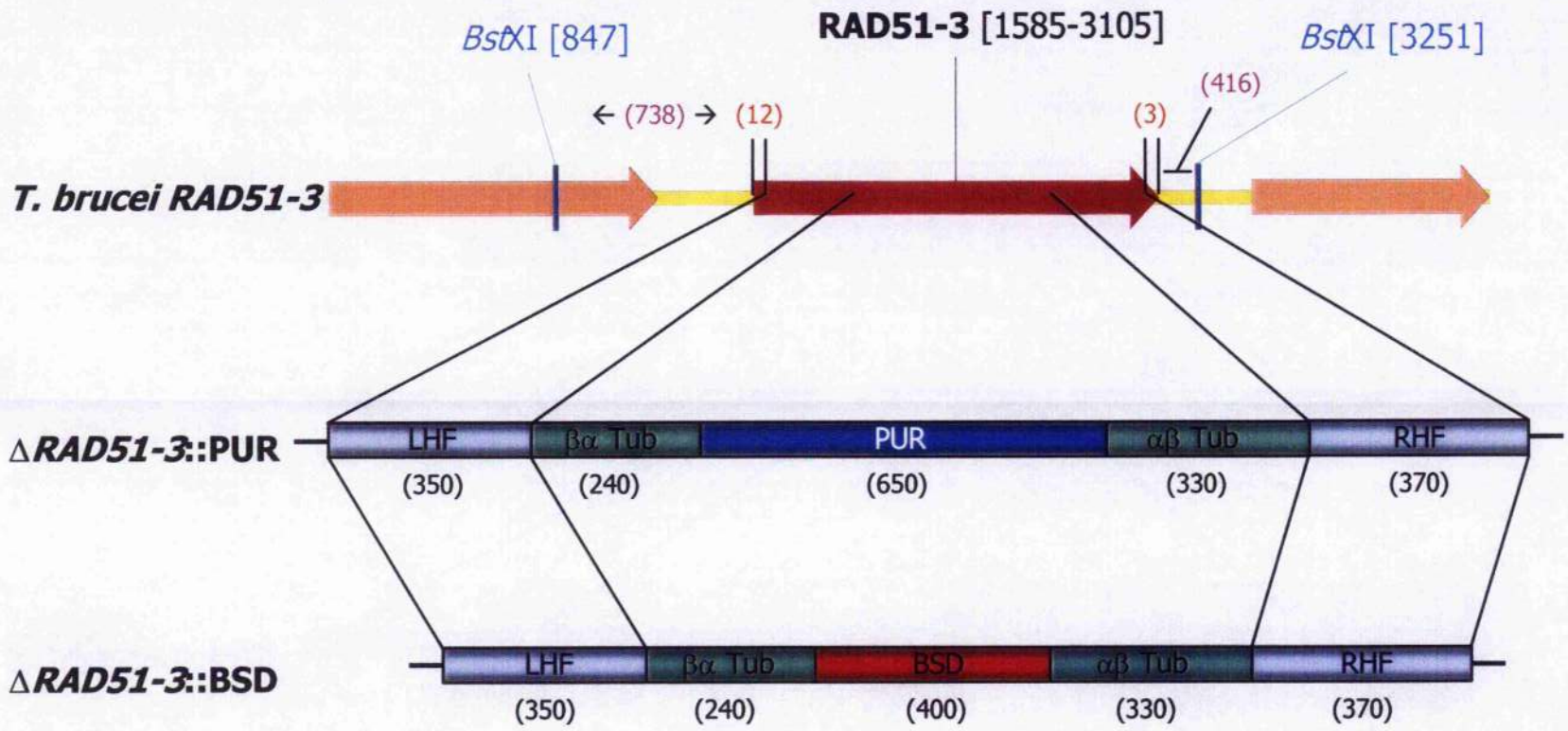


Figure 4.7: Disruption of the *RAD51-3* open reading frame. A diagram depicting disruption of both copies of the *RAD51-3* open reading frame and of the calculations carried out during the prediction of the *Bst*XI digestion fragments sizes. Figures in square brackets represent positions whereas figures in standard brackets are sizes (both shown in base pairs). Purple figures: distance between the open reading frame and the restriction site. Red figures: distance between the start of the open reading frame and the LHF or between the RHF and the end of the open reading frame. Calculations of fragment sizes were carried out as follows: Wild type: 3251 – 847 = 2404. Integration of Δ*RAD51-3*::*PUR*: 738 + 12 + 350 + 240 + 650 + 330 + 370 + 3 + 416 = 3109. Integration of Δ*RAD51-3*::*BSD*: 738 + 12 + 350 + 240 + 400 + 330 + 370 + 3 + 416 = 2859. See Figure 4.2 for Δ*RAD51-3*::*PUR* and Δ*RAD51-3*::*BSD* definitions and Figure 4.6 for *T. brucei RAD51-3* definitions.

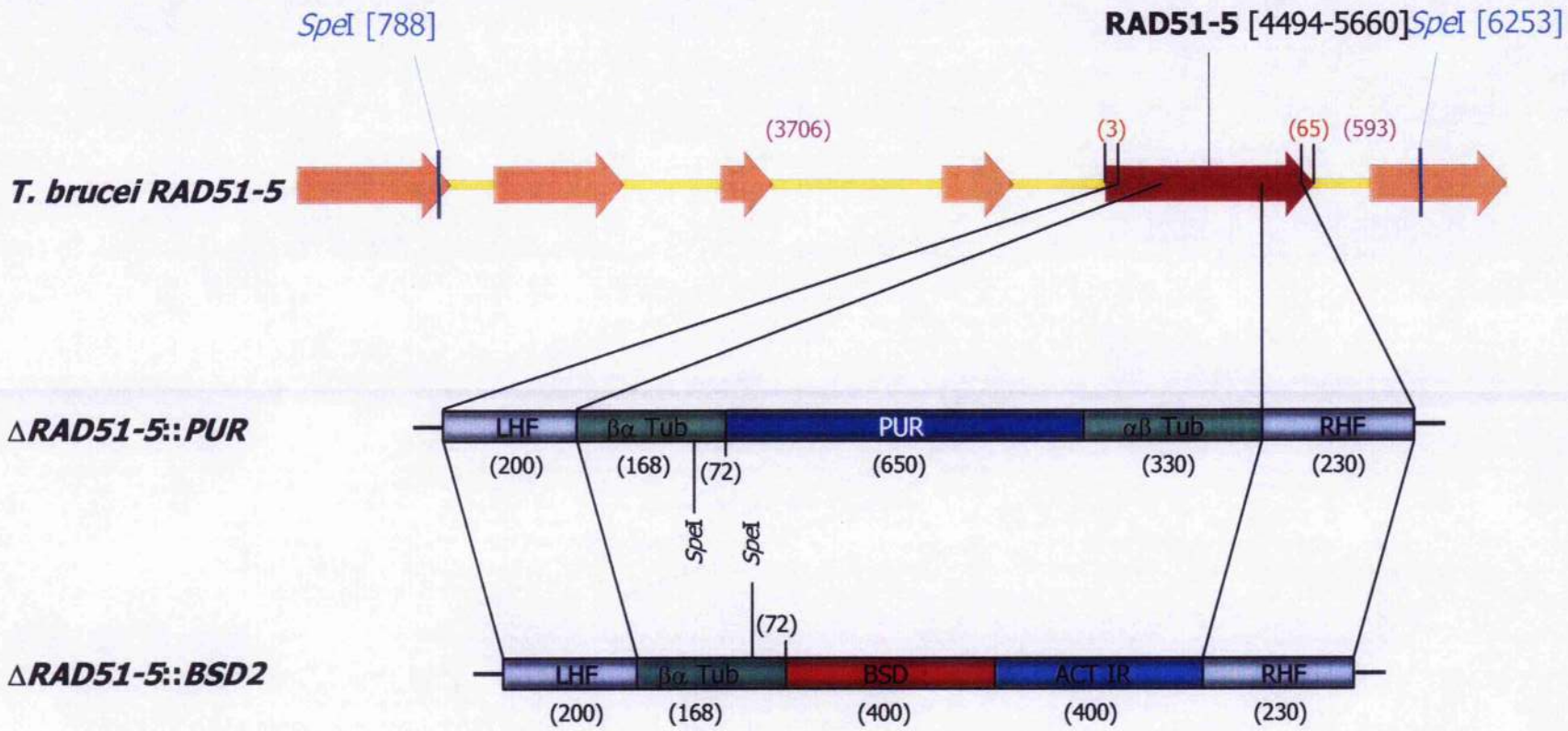


Figure 4.8: Disruption of the *RAD51-5* open reading frame. A diagram depicting disruption of both copies of the *RAD51-5* open reading frame and of the calculations carried out during the prediction of the *Spe*I digestion fragments sizes. Figures in square brackets represent positions whereas figures in standard brackets are sizes (both shown in base pairs). Purple figures: distance between the open reading frame and the restriction site. Red figures: distance between the start of the open reading frame and the LHF or between the RHF and the end of the open reading frame. Calculations of fragment sizes were carried out as follows: Wild type: 6253 – 788 = 5465. Integration of Δ*RAD51-5*::*PUR*: 3706 + 3 + 200 + 168 = 4077 & 72 + 650 + 330 + 230 + 65 + 593 = 1940. Integration of Δ*RAD51-5*::*BSD2*: 3706 + 3 + 200 + 168 = 4077 & 72 + 400 + 400 + 230 + 65 + 593 = 1760. See Figure 4.3 for Δ*RAD51-5*::*PUR* and Δ*RAD51-5*::*BSD2* definitions and Figure 4.6 for *T. brucei RAD51-5* definitions.

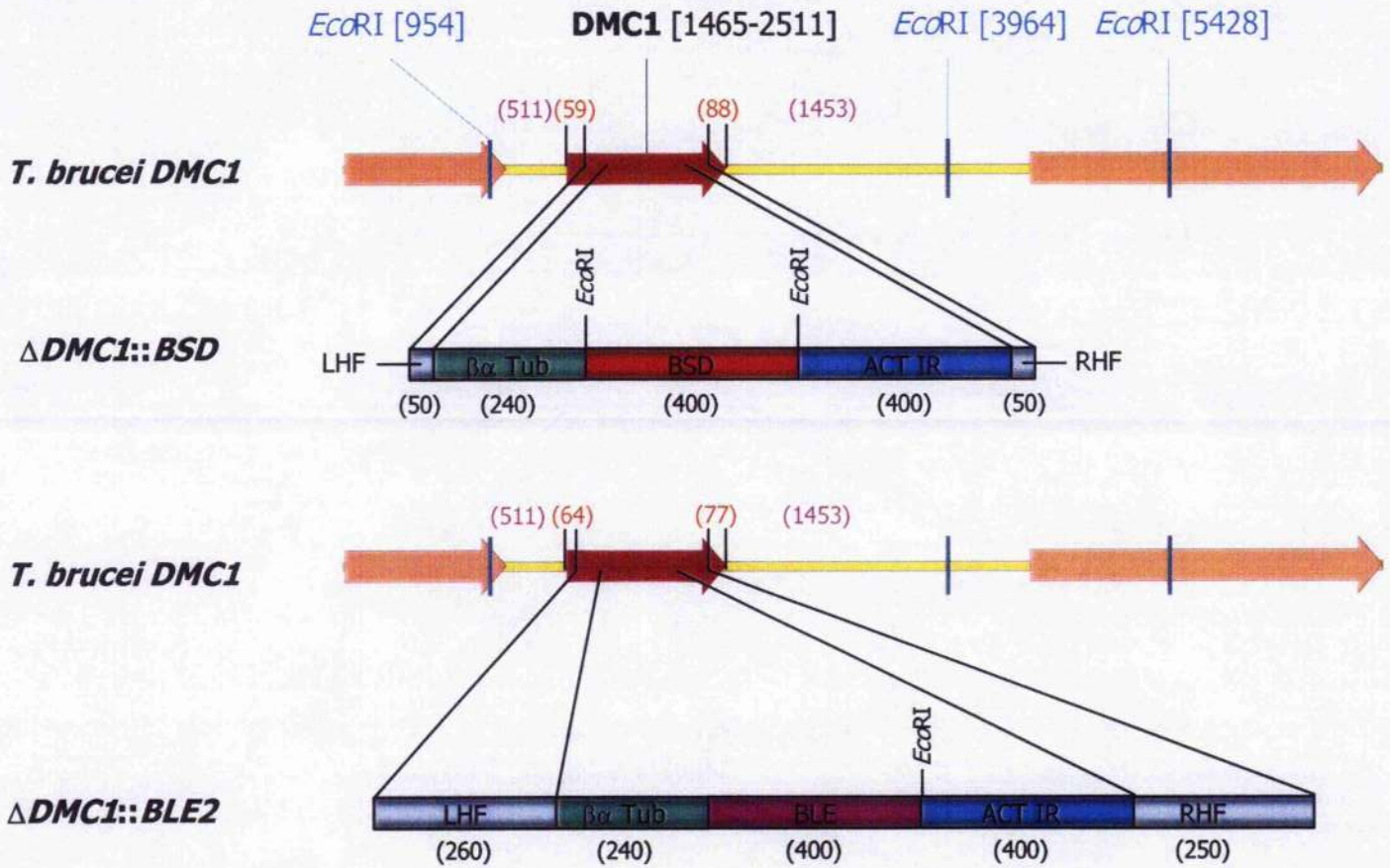


Figure 4.9: Disruption of the *DMC1* open reading frame. A diagram depicting disruption of both copies of the *DMC1* open reading frame and of the calculations carried out during the prediction of the *Eco*RI digestion fragments sizes. Figures in square brackets represent positions whereas figures in standard brackets are sizes (both shown in base pairs). Purple figures: distance between the open reading frame and the restriction site. Red figures: distance between the start of the open reading frame and the LHF or between the RHF and the end of the open reading frame. Calculations of fragment sizes were carried out as follows: Wild type: 3964 – 954 = 3010. Integration of Δ*DMC1*::*BSD*: 511 + 59 + 50 + 240 = 860 & 400 + 50 + 88 + 1453 = 1991. Integration of Δ*DMC1*::*BLE2*: 511 + 64 + 260 + 240 + 400 = 1475 & 400 + 250 + 77 + 1453 = 2180. See Figure 4.5 for Δ*DMC1*::*BSD* and Δ*DMC1*::*BLE2* definitions and Figure 4.6 for *T. brucei DMC1* definitions.

Southern analysis of the putative *RAD51-3* and *DMC1* mutants match the predicted fragment sizes (Fig. 4.10), indicating that the constructs have integrated into the correct locus, creating heterozygous and homozygous mutants for each gene. The blots were probed with the entire open reading frame and no band is discernable beyond the expected fragments, indicating that no intact open reading frame remains in the homozygous mutants. The fragments observed in the *RAD51-5* mutants do not, however, match those predicted (Fig 4.10); the wild-type undisrupted fragment and the larger fragment(s) in the heterozygous and homozygous alleles are smaller than those predicted, but the smaller fragments in the heterozygous and homozygous alleles match the predictions. This is most likely explained by an *Spe*I site that has been mis-sequenced upstream of *RAD51-5* in the *T. brucei* genome database or a base difference between the 927 *T. brucei* strain sequenced and the 3174.2 strain. This *Spe*I site must be approximately 3.2 kb upstream of *RAD51-5* to give the smaller fragment sizes that were observed. Despite this, Southern analysis suggests that the heterozygous and homozygous mutants for *RAD51-5* have been disrupted as expected. In each case, the disruption of one copy of the gene can be seen in the heterozygous cell lines and no wild type *RAD51-5* allele is retained in the homozygous cell lines.

4.2.6 Confirmation of the generation of mutants by Reverse Transcriptase-PCR

To confirm further the generation of homozygous mutants, Reverse Transcriptase–PCR (RT-PCR) was carried out. Total RNA was prepared from 25 ml of each cell line grown to approximately 2 x 10^6 cells.ml^{-1} (RNeasy kit, Invitrogen) of which 2µg was DNase I treated (DNase Amplification Grade, Life Technologies) before cDNA was generated using random oligonucleotides and reverse transcriptase (Superscript first strand synthesis system, Life Technologies). For each cDNA prepared, a reverse transcriptase minus reaction was carried out to control for any genomic DNA contamination that may be present.

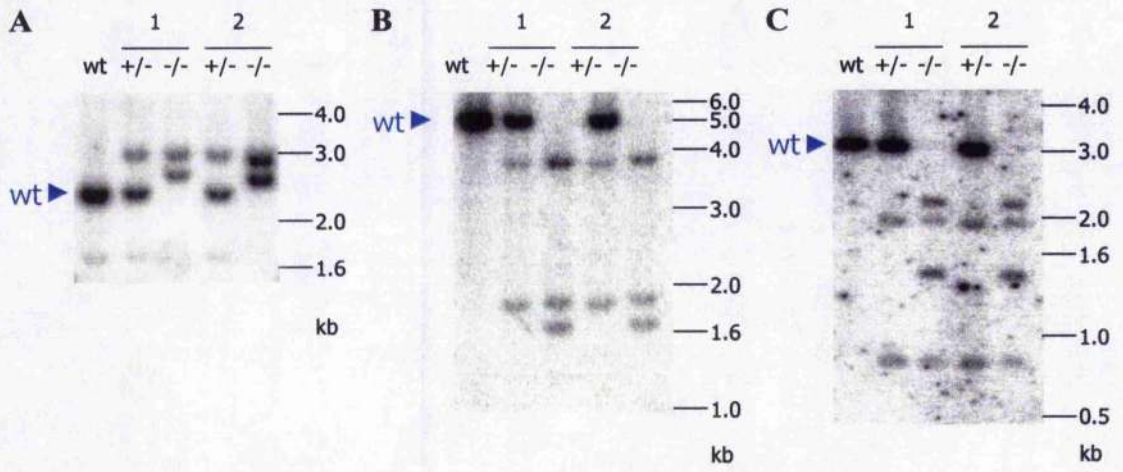


Figure 4.10: Southern analysis of the *RAD51-3*, *RAD51-5* and *DMC1* mutant cell lines. (A) *RAD51-3* cell lines digested with *Bst*XI, (B) *RAD51-5* cell lines digested with *Spe*I and (C) *DMC1* cell lines digested with *Eco*RI. 5 µg of genomic DNA of each cell line was restriction digested for 6 hrs before being run out on a 0.8% agarose gel. The DNA was Southern blotted before being probed with the entire open reading frame. The two independent heterozygous mutants are indicated by +/- 1 and 2 and the homozygous mutants are indicated by -/- 1 and 2. WT refers to genomic DNA from the untransformed 3174.2 cells. The size of the fragment that corresponds to the undisrupted gene is shown (wt), as are size markers.

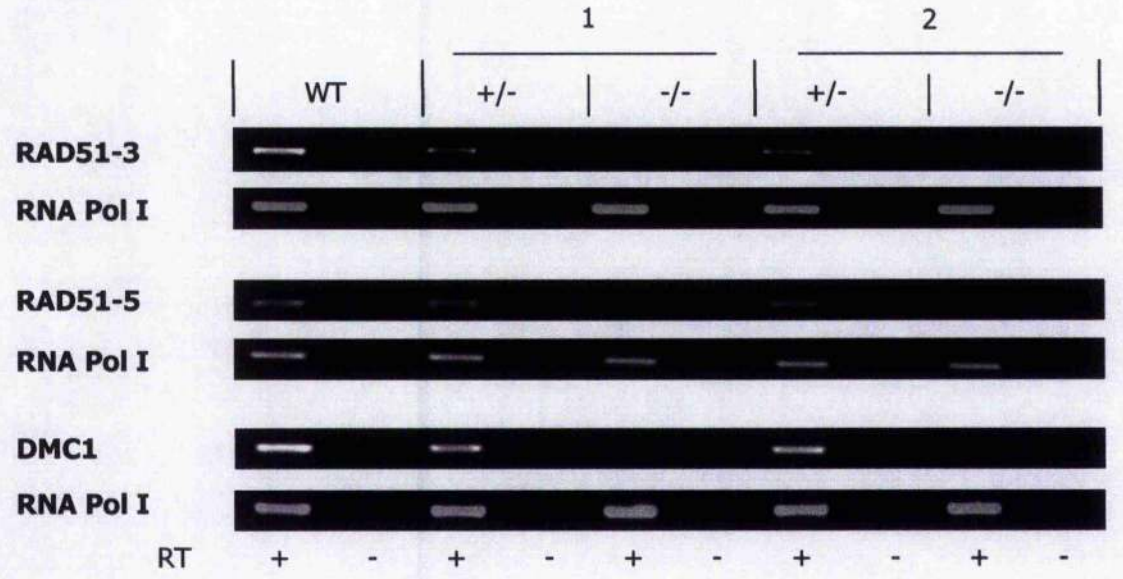


Figure 4.11: Confirmation of the generation of mutants by RT-PCR. RT-PCR was carried out on cDNA generated form total RNA. RNA Polymerase I-specific primers were used to show the generation of intact cDNA. Primers specific for *RAD51-3*, *RAD51-5* or *DMC1* were used to assay expression of the genes in wild type (WT) cells, heterozygous mutants (+/-) and homozygous mutants (-/-).

RT-PCR was carried out using gene-specific primers to assay for the presence or absence of intact RNA corresponding to each gene in the mutants. For *RAD51-3*, the primers *RAD51-3*-KO5′ and *RAD51-3*-KO3′ (Appendix 1) were used which should give a product of 669 bp (Fig 4.11). For *RAD51-5*, the primers *RAD51-5*- LHF-For and *RAD51-5*-U2 (Appendix 1) were used which should give a product of 435 bp (Fig. 4.11). For *DMC1*, the primers *DMC1*-LHF-For and *DMC1*-U1 (Appendix 1) were used which should give a product of 340 bp (Fig. 4.11). In each case, a specific product of the appropriate size was generated in the wild type and heterozygous cell lines. However, disruption of both alleles of the gene in the homozygous mutants resulted in no PCR product being generated, confirming that intact RNA for each gene is not transcribed in any of the respective homozygous mutants (Fig. 4.11).

4.3 Analysis of growth

The growth rate of the *RAD51-3*, *RAD51-5* and *DMC1* cell lines were analysed for two reasons. Firstly, any observed defect in growth potentially relates to the importance of that gene in the completion of cell division, at least in the particular life cycle stage being studied. For example, it has been observed previously that *rad51-/-* mutants have a significantly increased population doubling time (McCulloch and Barry, 1999) relative to wild type cells, as do *mre11-/-* mutants (Robinson *et al.*, 2002). In contrast, other genes involved in DNA repair reactions such as *KU70* (Conway *et al.*, 2002a), *MHS2* and *MLH1* (Bell and McCulloch, 2003) do not show such growth impairment. Secondly, any *in vitro* or *in vivo* growth defect that we observe will need to be taken into account when carrying out and interpreting further assays, such as VSG switching.

4.3.1 Analysis of *in vitro* growth

in vitro growth analysis was carried out on each of the heterozygous and homozygous cell lines for each gene and compared with wild type 3174.2 cells. To do this, 5 ml cultures were inoculated at cell densities 5 x 10^4 or 1 x 10^5 cells.ml^{-1} and counted using a haemocytometer (bright-line, Sigma) at 24, 48 and 72 hours subsequently. Two or three repetitions of each cell line were carried out, the results plotted on a semi-

logarithmic scale (Fig. 4.12) and the doubling times for the heterozygous and homozygous mutants calculated and compared to that of the wild type cells (Table 4.5).

	Cell line				
Gene	Wild-type	1 +/-	1 -/-	2 +/-	2 -/-
RAD51-3	8.6	8.4	15.7	8.7	15.0
RAD51-5	7.8	8.0	12.8	7.6	12.6
DMC1	8.0	8.1	8.3	8.0	8.0

Table 4.5: *in vitro* population doubling times for *RAD51-3*, *RAD51-5* and *DMC1* mutants. The average doubling time for each of the independent heterozygous (+/-) and homozygous (-/-) mutants is displayed in hours. The table also includes the average doubling time for 3174.2 wild type (WT) cells.

From the growth curves it can clearly be seen that disruption of one allele of any of the genes, *RAD51-3*, *RAD51-5* or *DMC1*, had no effect on growth, since all the heterozygotes grew at the same rate as 3174.2 wild type cells. Furthermore, the *dmc1* -/- cells, showed no growth defect when compared with either the wild type or *DMC1* +/- cells. In contrast, *rad51-3*-/- and *rad51-5*-/- showed marked increases in doubling time relative to 3174.2 wild type cells and their respective heterozygous mutants.

The calculated doubling time for the *rad51-5*-/- cells was approximately 12.7 hrs, suggesting the mutant has caused a similar growth defect to that observed in *rad51*-/- mutants, which double in approximately 11 hrs (McCulloch and Barry, 1999). The *rad51-3*-/- mutants appeared to have an even greater growth defect, doubling in approximately 15.3 hours.

4.3.2 Analysis of *in vivo* growth

in vivo growth was analysed to determine if the growth defects that we observed *in vitro* are exacerbated or alleviated during growth in mice. This is particularly important when analysing the effect of loss of the *RAD51*-like genes on VSG switching, as the number of cell divisions in the mouse bloodstream needs to be factored in when calculating the switching rate (see Section 4.8.1). *In vivo* growth counts were carried out for each of the *RAD51-3* and *RAD51-5* mutant cell lines and 3174.2 wild-type cells. The *DMC1* mutant cell lines were not analysed.

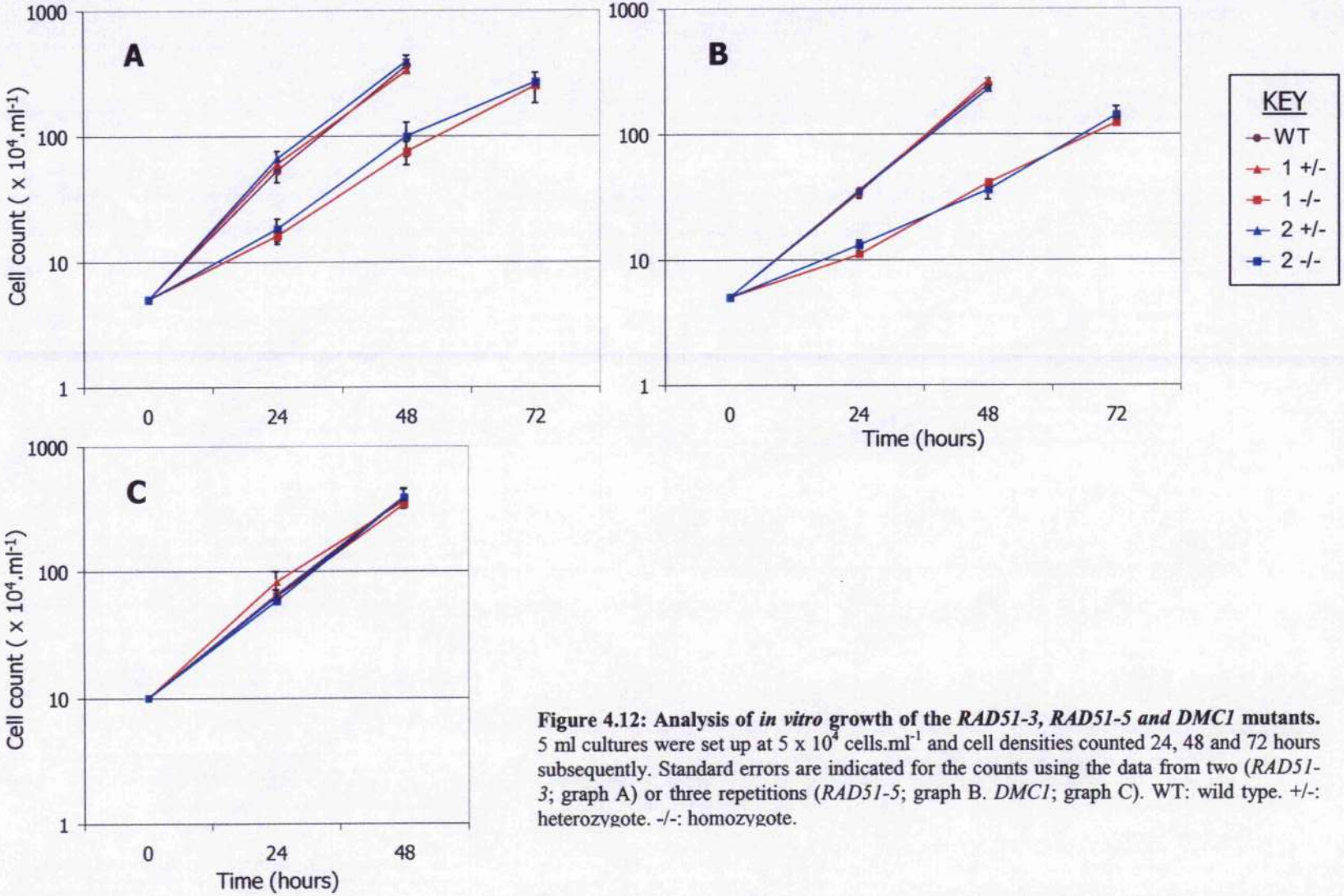


Figure 4.12: Analysis of *in vitro* growth of the *RAD51-3, RAD51-5 and DMC1* mutants. 5 ml cultures were set up at 5 x 10^4 cells.ml^{-1} and cell densities counted 24, 48 and 72 hours subsequently. Standard errors are indicated for the counts using the data from two (*RAD51-3*; graph A) or three repetitions (*RAD51-5*; graph B. *DMC1*; graph C). WT: wild type. +/-: heterozygote. -/-: homozygote.

ICR mice were infected with 1 x 10^6 *T. brucei*, previously grown in culture, via interperitoneal injection, and the density of trypanosomes was then determined every 24 hours up to a maximum of 120 hours. To do this, a small volume of blood was removed from the tail of each mouse into heparin-coated capillary tubes (Hawksley) of which 1 µl samples were then diluted in 99 µl of 0.85% ammonium chloride. This lyses preferentially the red blood cells, allowing the *T. brucei* cells to be counted using a haemocytometer (bright-line, Sigma). The results were plotted on semi-logarithmic graphs (Fig. 4.13) and the doubling time for the heterozygous and homozygous mutants calculated and compared to that of the 3174.2 wild type (Table 4.6). Note that for each cell line, only a single mouse infection was performed, meaning that each set of data represents a single experiment.

	Cell line				
Gene	Wild-type	1 +/-	1 -/-	2 +/-	2 -/-
RAD51-3	5.5	5.5	8.7	5.5	7.8
RAD51-5	5.5	5.3	9.5	5.1	9.8

Table 4.6: *in vivo* doubling times for *RAD51-3* and *RAD51-5* mutants. The doubling time for each of the heterozygous (+/-) and homozygous (-/-) mutants is displayed in hours. The table also includes the doubling time for 3174.2 wild type (WT) cells.

The population doubling times for all of the cell lines appear to be quicker *in vivo* when compared with the *in vitro* data. The wild type and heterozygous cell lines all appear to double in 5-5.5 hours, compared with 7.6-8.7 hours seen *in vitro*. The *rad51-3-/-* mutants appeared to have a slightly less severe growth defect in *vivo* compared with the data derived from growth in culture. The average population doubling time *in vivo* of the two lines (8.25 hrs) is a 1.5-fold increase relative to the wild type and *RAD51-3+/-* mutants (assuming an average doubling time of 5.5 hrs). In contrast, *in vitro* the *rad51-3-/-* mutants doubled, on average, every 15.35 hours, which represents a 1.8-fold increase relative to the wild type and *RAD51-3 +/-* mutants (average of 8.57 hrs). Conversely, the *rad51-5-/-* mutants appear to have a slightly more pronounced growth defect *in vivo*. The average population doubling time *in vivo* of the two lines (9.65 hrs) is a 1.8 fold increase relative to the wild type and *RAD51-5+/-* mutants (assuming an average doubling time of 5.3 hrs). *in vitro* the *rad51-5-/-* mutants doubled, on average, every 12.7 hours, which represents a 1.6-fold increase relative to the wild type and *RAD51-5 +/-* mutants (average of 7.8 hrs). The *in vitro* population doubling times will be used during the calculations of the VSG switching frequency (see Section 4.8.1), as

more repetitions have been carried out and it is therefore likely that the data is more accurate.

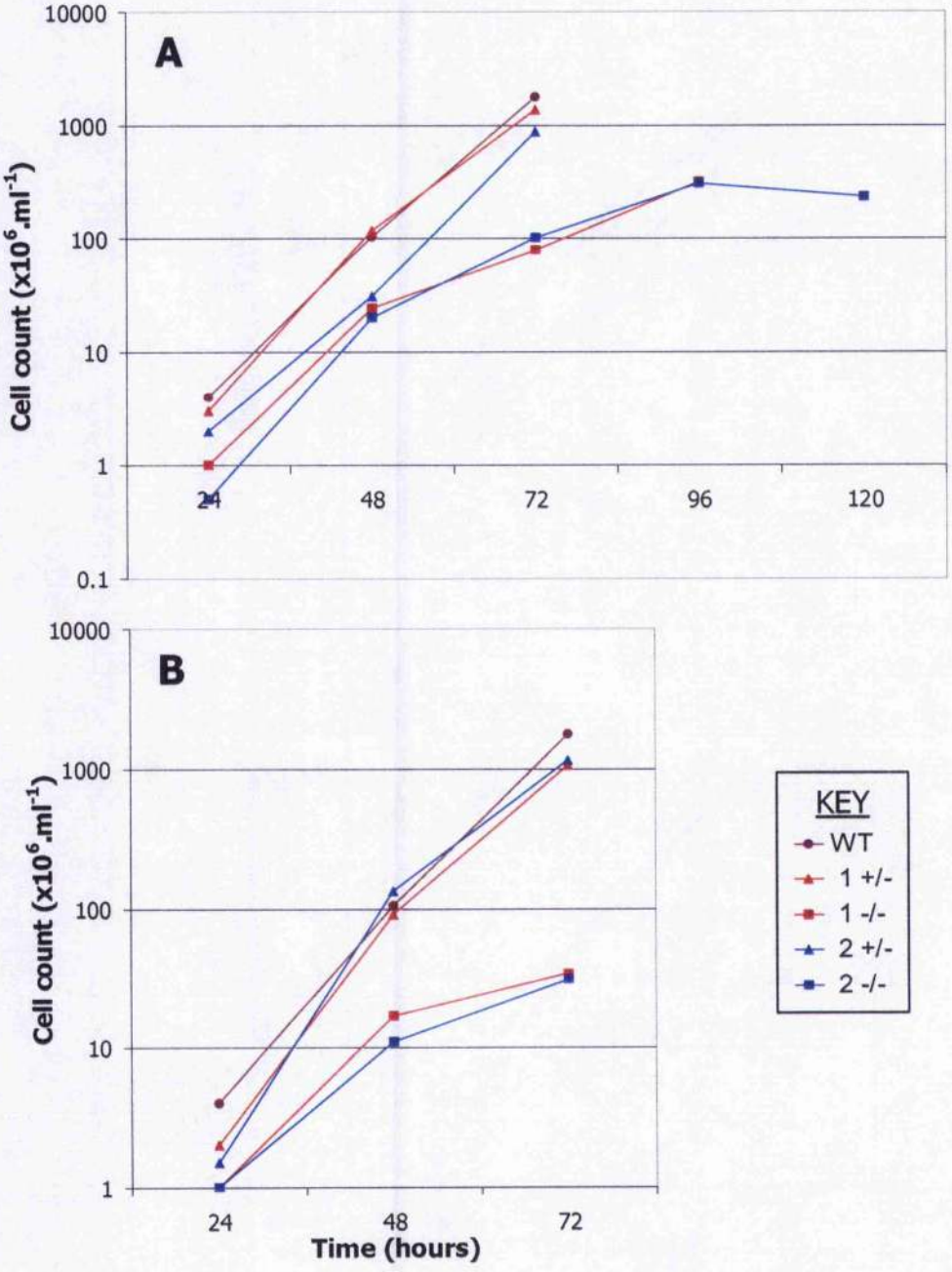


Figure 4.13: Analysis of the growth of *RAD51-3* and *RAD51-5* mutants *in vivo*. Mice were infected with 1 x 10^6 trypanosomes, previously grown in culture, *via* interperitoneal injection. The density of trypanosomes was then recorded every 24 hours up to a maximum of 120 hours. One experiment was carried out for each of the heterozygous (+/-) and homozygous (-/-) mutants and wild type (WT) cells and the results plotted on semi-logarithmic scale graphs, represent. A: *RAD51-3*. B: *RAD51-5*.

4.4 Analysis of the cell cycle in the *RAD51-3*, *RAD51-5* and *DMC1* mutants

Given that mutation of *RAD51-3* and *RAD51-5* leads to increased population doubling, we wished to determine if the disruption had affected the cell cycle. The increased population doubling could be explained either by each mutant cell taking longer to complete the cell cycle, or increased numbers of cells that fail to complete cell division in the mutants. Kinetoplastids have a unique feature that allows for definition of the cell cycle point of individual cells by DNA stains. This is because the DNA contained within the mitochondrion (kinetoplast) and the nucleus is replicated in distinct synthesis phases. Kinetoplast replication begins before, and is shorter that, nuclear replication, resulting in the completion of kinetoplast division before the nuclear DNA enters mitosis (McKean, 2003). Cells in G1 phase of the cell cycle contain 1 nucleus and 1 kinetoplast (1N 1K). Kinetoplast division, which results in cells with 1 nucleus and 2 kinetoplasts (1N 2K), occurs prior to mitosis, during which time the cells have 2 nuclei and 2 kinetoplasts (2N 2K). Completion of cell division forms two 1N1K cells in G1 phase. This means that even asynchronous cultures can have their cell cycle stage accurately determined (McKean, 2003; Fig 4.14).

To examine the cell cycle, we used 4′,6-Diamidino-2-phenyindole (DAPI) stain, which binds to DNA and fluoresces under UV light, to analyse the DNA content of fixed *T. brucei* cells (see Section 2.13.1). Comparison of the homozygous mutant cells with the heterozygous and wild-type cells will, therefore, determine if any mutant has a affected the cell cycle. The *RAD51-3*, *RAD51-5* and *DMC1* strains were grown *in vitro* to a density of approximately 1 x 10^6 cells.ml^{-1}. 1 ml of this was centrifuged, the cells washed twice with PBS and then resuspended in 1 ml of PBS. 10 µl samples were then spotted onto multi-spot microscope slides (C.A.Hendley Ltd.) and allowed to air-dry. The trypanosomes were then fixed by soaking in methanol for 5 minutes at room temperature and again allowed to air dry before vectashield with DAPI (Vector Laboratories Inc.) was added, a coverslip positioned and the slide sealed with clear nail varnish. Differential interference contrast (DIC) was then used to visualise intact cells and UV used to visualise DAPI (see Section 2.13 for further details). 2 counts of approximately 100 cells for each cell line were carried out (one of which was independent and conducted blind by R. McCulloch). The counts were added and the percentages of cells with each discernable DNA content calculated (Figs. 4.15-17).

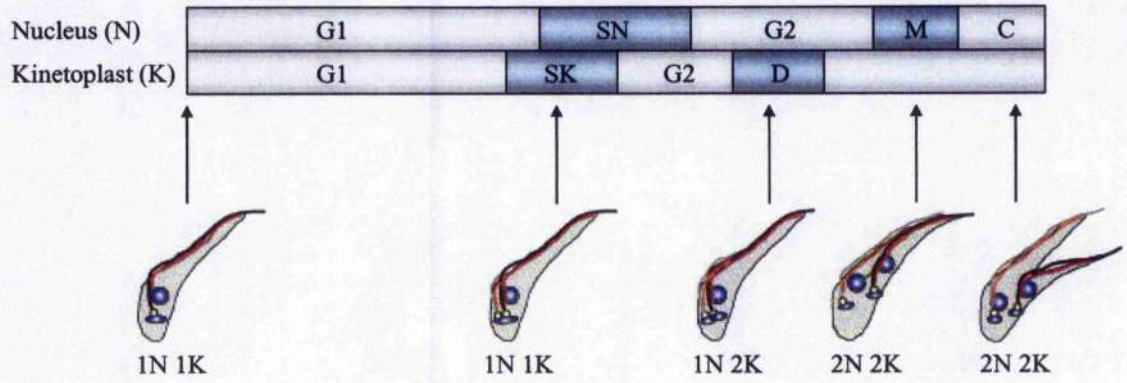


Figure 4.14: Cell cycle of *T. brucei*. The diagram shows the timing differences between the replication and division of the nucleus and kinetoplast during the cell cycle. Trypanosomes start the cell cycle containing 1 nucleus and 1 kinetoplast (1N 1K). Kinetoplast synthesis (SK) then begins, which is shorter than, and starts prior to, nuclear synthesis (SN), meaning that kinetoplast division (D), which results in cells that have 1 nucleus and 2 kinetoplasts (1N 2K), has completed before nuclear mitosis (M) starts. Nuclear mitosis, where cells have 2 nuclei and 2 kinetoplasts (2N 2K), occurs prior to cytokinesis (C) and marks the end of DNA replication and division and divides the cell to generate two progeny (1N 1K), which restart the cell cycle. Figure adapted from McKean (2003).

For all three of the *RAD51*-like genes their mutation appeared to have no effect on the cell cycle (Figs. 4.15-17). In all cases, the numbers of cells in each cell cycle stage were equivalent in the wild types, heterozygous mutants and homozygous mutants. The growth defect appears not, therefore, to be a consequence of any of the mutants stalling at a discernible stage during the cell cycle. The same appears to be the case for *rad51-/- T. brucei*. The cells described as 'other' in this assay are those that have a DNA content that does not fall into one of the expected categories. This can take the form of too many nuclei or kinetoplasts, or an absence of one or both. These cells can arise due to defects during DNA replication, or occur as a result of inappropriate segregation of the nuclei and kinetoplasts during cell division. These cells appear to account for approximately 2-4% of a 3174.2 wild type population and appear not to increase in number as a result of disruption of *RAD51-3*, *RAD51-5* or *DMC1*.

	Cell cycle stage			
Cell line	**1N1K**	**1N2K**	**2N2K**	**Other**
Wild type	155	25	15	5
	131	23	9	8
Total	**286**	**48**	**24**	**13**
1 +/-	157	38	17	5
	122	22	12	9
Total	**279**	**60**	**29**	**14**
1 -/-	168	28	12	7
	96	17	3	4
Total	**264**	**45**	**15**	**11**
2 +/-	160	24	26	5
	101	16	17	6
Total	**261**	**40**	**43**	**11**
2 -/-	144	37	19	8
	117	28	8	9
Total	**261**	**65**	**27**	**17**
RAD51-/-	190	19	13	3
	145	17	8	5
Total	**335**	**36**	**21**	**8**

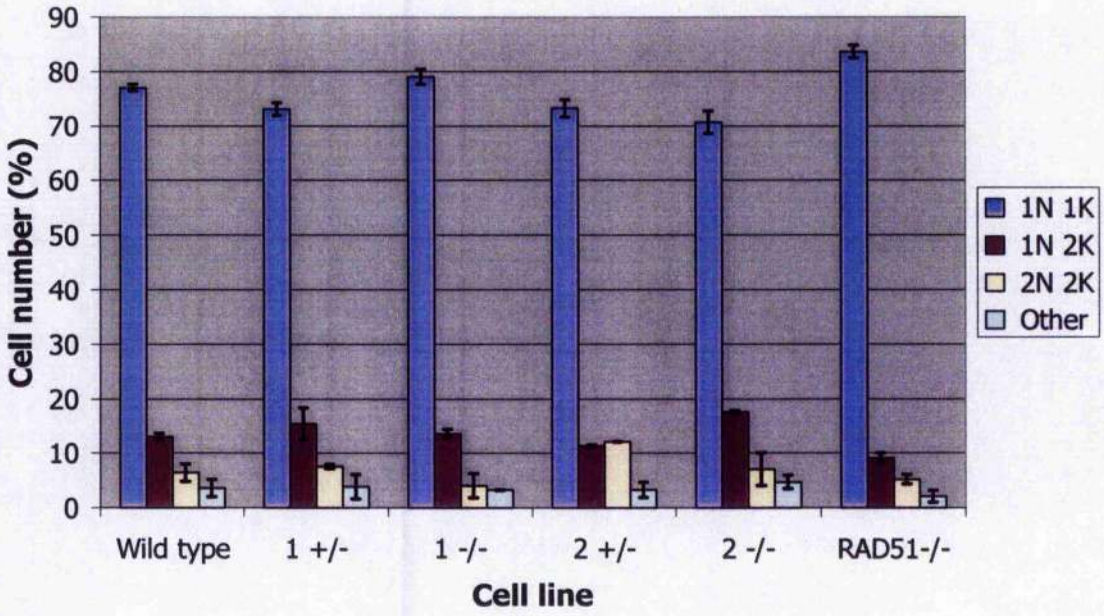


Figure 4.15: DAPI analysis of the *RAD51-3* mutants. The DNA content of the 2 *RAD51-3* heterozygous (+/-) and homozygous (-/-) mutant cell lines was visualised by DAPI and compared with *rad51* homozygous (-/-) mutants and wild type (WT) cells. The numbers of cells counted, in two experiments for each cell line, are tabulated (1N1K: 1 nucleus, 1 kinetoplast; 1N2K: 1 nucleus, 2 kinetoplast; 2N2K: 2 nuclei, 2 kinetoplast; Other: cells that did not fall into the previous categories). The average percentages of the two experiments are graphed, including standard deviations generated from the two data sets.

	Cell cycle stage			
Cell line	**1N1K**	**1N2K**	**2N2K**	**Other**
Wild type	83	20	7	4
	88	20	2	1
Total	**171**	**40**	**9**	**5**
1 +/-	79	20	12	4
	74	32	2	2
Total	**153**	**52**	**14**	**6**
1 -/-	65	31	14	3
	75	18	4	2
Total	**140**	**49**	**18**	**5**
2 +/-	88	15	6	3
	76	16	2	5
Total	**164**	**31**	**8**	**8**
2 -/-	73	32	8	2
	76	39	11	3
Total	**149**	**71**	**19**	**5**

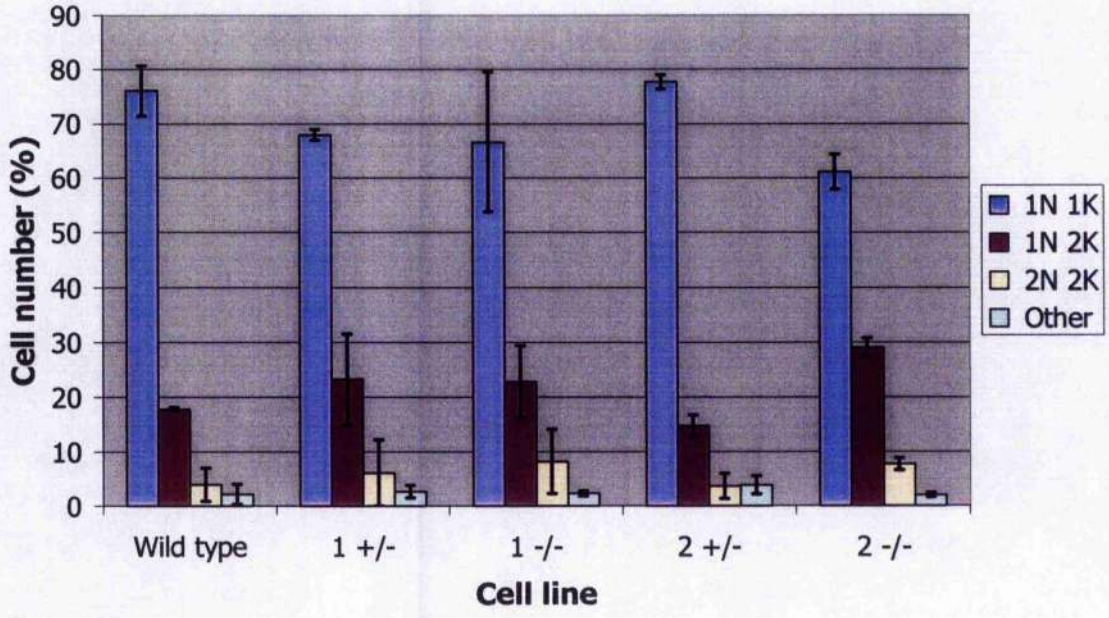


Figure 4.16: DAPI analysis of the *RAD51-5* mutants. The DNA content of the 2 *RAD51-5* heterozygous (+/-) and homozygous (-/-) mutant cell lines was visualised by DAPI and compared with wild type (WT) cells. The numbers of cells counted, in two experiments for each cell line, are tabulated (1N1K: 1 nucleus, 1 kinetoplast; 1N2K: 1 nucleus, 2 kinetoplast; 2N2K: 2 nuclei, 2 kinetoplast; Other: cells that did not fall into the previous categories). The average percentages of the two experiments are graphed, including standard deviations generated from the two data sets.

	Cell cycle stage			
Cell line	**1N1K**	**1N2K**	**2N2K**	**Other**
Wild type	85	15	5	4
	46	9	2	3
Total	**131**	**24**	**7**	**7**
1 +/-	66	20	14	6
	38	10	4	3
Total	**104**	**30**	**18**	**9**
1 -/-	70	24	6	5
	51	10	6	4
Total	**121**	**34**	**12**	**9**
2 +/-	82	13	6	7
	47	17	5	3
Total	**129**	**30**	**11**	**10**
2 -/-	88	20	9	5
	42	10	5	2
Total	**130**	**30**	**14**	**7**

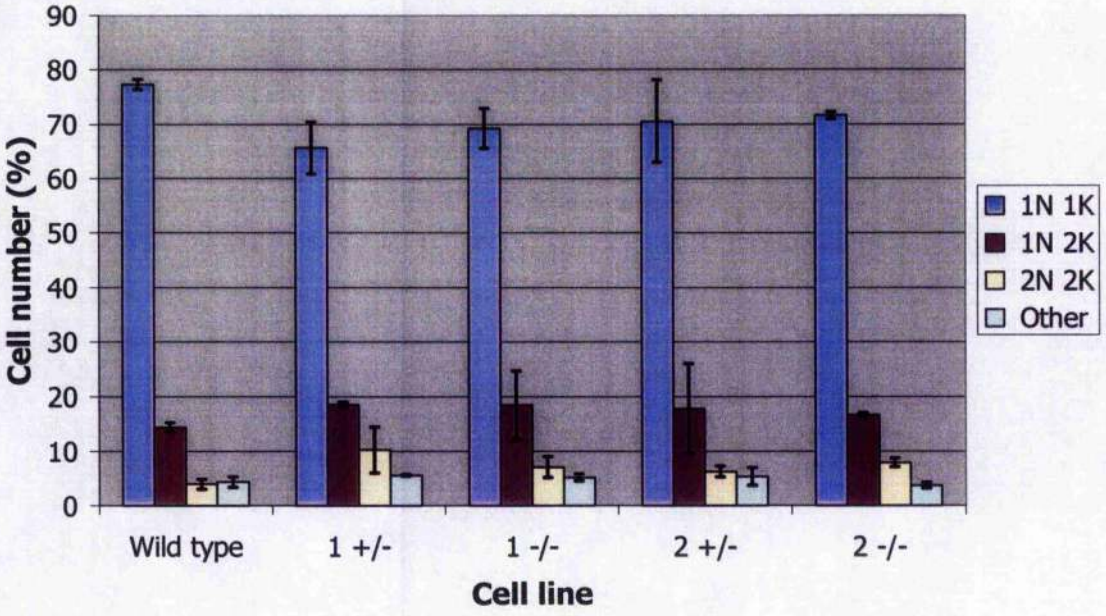


Figure 4.17: DAPI analysis of the *DMC1* mutants. The DNA content of the 2 *DMC1* heterozygous (+/-) and homozygous (-/-) mutant cell lines was visualised by DAPI and compared with wild type (WT) cells. The numbers of cells counted, in two experiments for each cell line, are tabulated (1N1K: 1 nucleus, 1 kinetoplast; 1N2K: 1 nucleus, 2 kinetoplast; 2N2K: 2 nuclei, 2 kinetoplast; Other: cells that did not fall into the previous categories). The average percentages of the two experiments are graphed, including standard deviations generated from the two data sets.

4.5 Analysis of DNA damage sensitivity in the *RAD51-3*, *RAD51-5* and *DMC1* mutants

To analyse whether or not the *RAD51*-like genes have a role in DNA repair a DNA damage assay was carried out by plating out cells in increasing concentrations of methyl methane sulphonate (MMS). MMS is a methylation agent that is capable of modifying DNA at a number of different sites, resulting in the generation of lethal and mutagenic lesions (Sedgwick, 2004). Each strain, previously grown to a density of 1 x 10^6 cells.ml^{-1}, was plated at one cell per well in five 96 well plates, each plate containing a different concentration of MMS: 0, 0.0001, 0.0002, 0.0003 and 0.0004%. Four repetitions for each strain were carried out. The number of wells containing a trypanosome population after 20 days of growth for each strain at each MMS concentration was counted and compared with the number of wells to grow through on the 0% MMS control plate for that strain. The number of wells to grow through on the no MMS control plates was averaged and counted as 100%; the results for that strain on the MMS-containing plates was then expressed relative to 100%, therefore removing any error due to plating efficiency.

The results observed for the *RAD51-3* (Fig 4.18) and *RAD51-5* mutants (Fig. 4.19) were very similar, with the disruption of one allele of either gene having no effect on the sensitivity to MMS when compared to wild-type cells. The homozygous mutants, however, displayed a marked increase in sensitivity, since no (or very limited) cell growth occurred at 0.0003% MMS, whereas the wild type and heterozygous mutants showed between 60 and 80% survival rates. Moreover at 0.0001% and 0.0002% MMS, here the wild type and heterozygous mutants showed between 80-100% survival, the homozygous mutants showed between 40-80% and 10-40% survival respectively. This result is very similar to that observed in *rad51-/-* cells (McCulloch and Barry, 1999) suggesting that both *RAD51-3* and *RAD51-5* have functions in DNA damage repair that are comparable with *RAD51*.

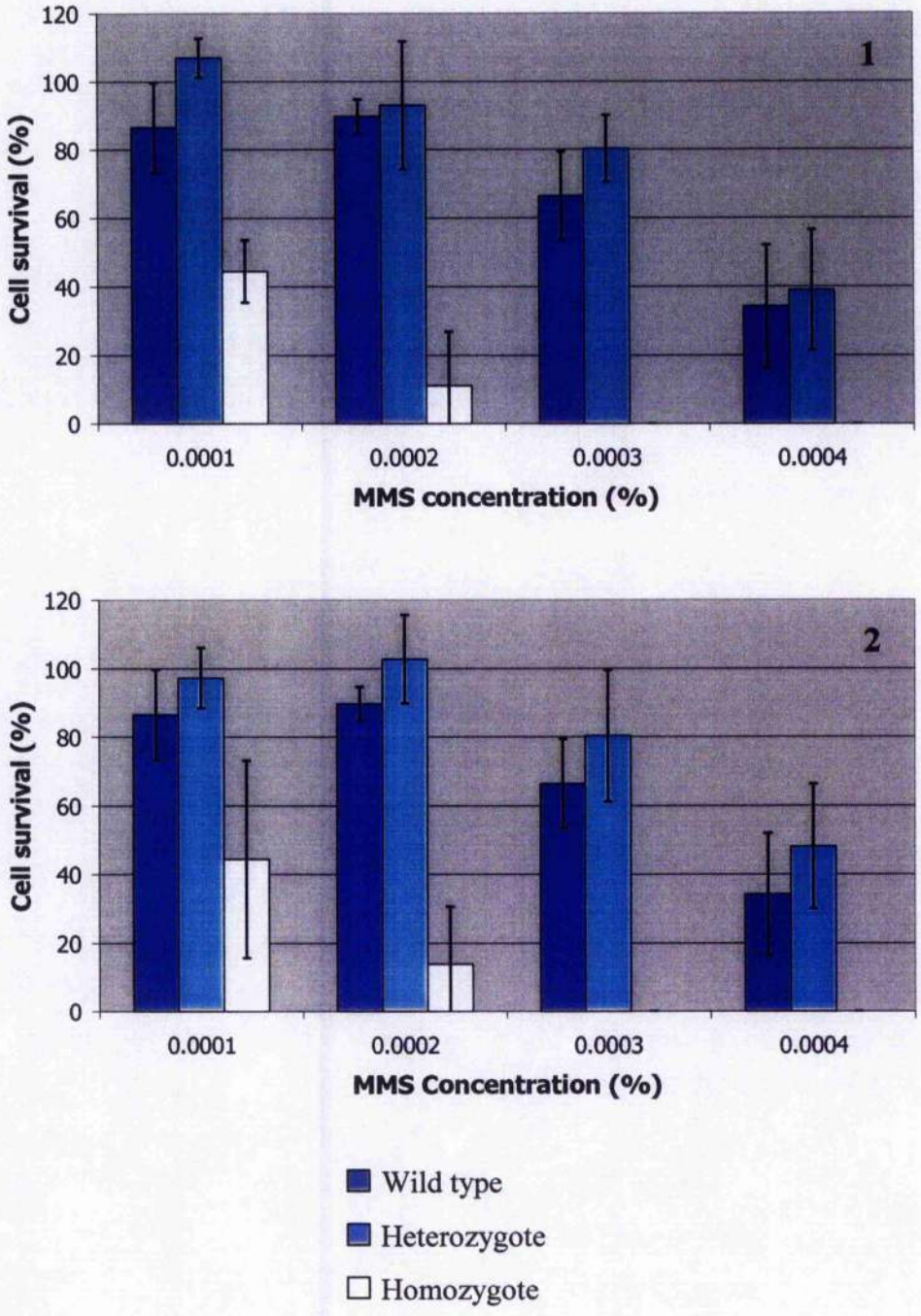


Figure 4.18: Analysis of DNA damage sensitivity in the *RAD51-3* mutants. Each strain was plated at one cell.well^{-1} in five 96 well plates, each plate containing a different concentration of MMS: 0, 0.0001, 0.0002, 0.0003 and 0.0004%. Four repetitions for each strain were carried out. The average number of wells to grow through for each strain at each MMS concentration was calculated and compared to the average number of wells to grow through on the 0% MMS control plate for that strain. Standard deviation is indicated and the data is presented for the two independent (1 and 2) *RAD51-3* heterozygous and homozygous mutants compared to 3174.2 wild type cells.

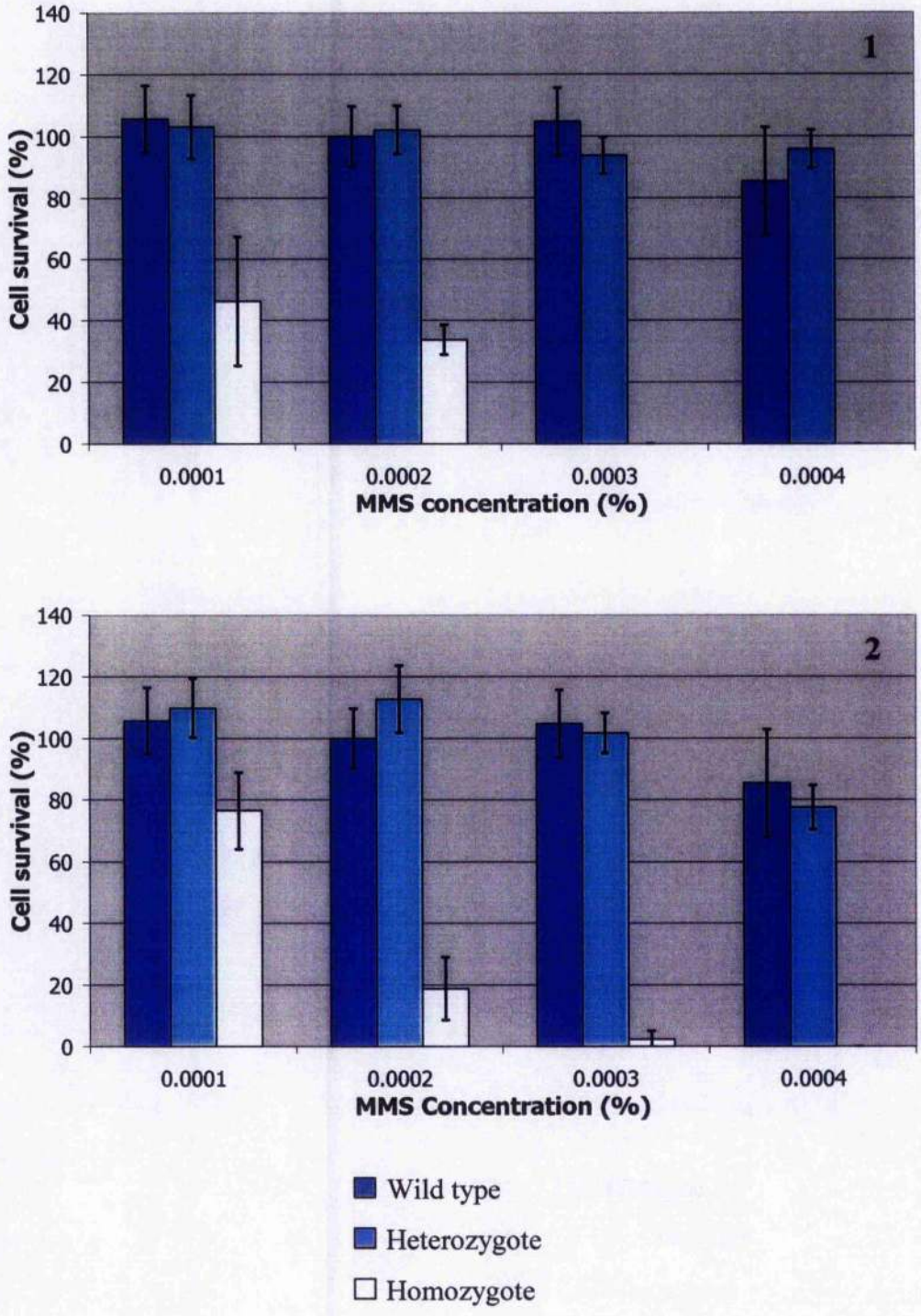


Figure 4.19: Analysis of DNA damage sensitivity in the *RAD51-5* mutants. Each strain was plated at one cell.well^{-1} in five 96 well plates, each plate containing a different concentration of MMS: 0, 0.0001, 0.0002, 0.0003 and 0.0004%. Four repetitions for each strain were carried out. The average number of wells to grow through for each strain at each MMS concentration was calculated and compared to the average number of wells to grow through on the 0% MMS control plate for that strain. Standard deviation is indicated and the data is presented for the two independent (1 and 2) *RAD51-5* heterozygous and homozygous mutants compared to 3174.2 wild type cells.

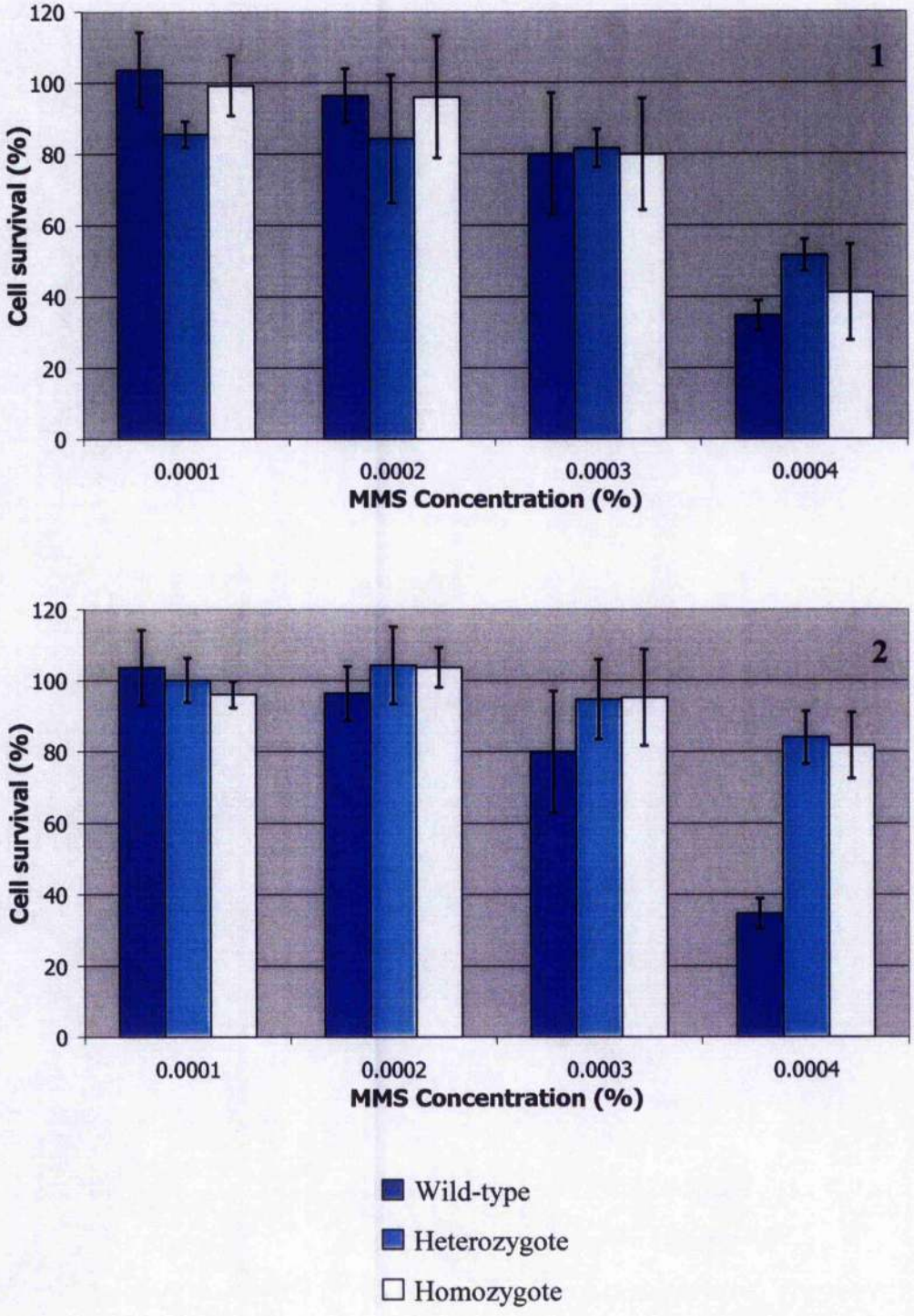


Figure 4.20: Analysis of DNA damage sensitivity in the *DMC1* mutants. Each strain was plated at one cell.well^{-1} in five 96 well plates, each plate containing a different concentration of MMS: 0, 0.0001, 0.0002, 0.0003 and 0.0004%. Four repetitions for each strain were carried out. The average number of wells to grow through for each strain at each MMS concentration was calculated and compared to the average number of wells to grow through on the 0% MMS control plate for that strain. Standard deviation is indicated and the data is presented for the two independent (1 and 2) *DMC1* heterozygous and homozygous mutants compared to 3174.2 wild type cells.

DMC1, on the other hand, appears to have no detectable role in MMS damage repair, since neither of the heterozygous or homozygous mutant lines generated showed an increase in DNA damage sensitivity (Fig 4.20). *DMC1*+/- 2 and *dmc1*-/- 2 (the later mutant derived from the former) appeared to show increased resistance to MMS, since at 0.0004% MMS (and perhaps at 0.0003% MMS) they showed greater survival than the wild type cells. This finding is not a result of the loss of *DMC1*, however, as no further increased resistance to MMS following the disruption of the second copy of the gene is observed. During the generation of *DMC1*+/- 2, we hypothesise that a mutation must have arisen resulting in resistance to MMS. This is not unprecedented, as mutants with increased resistance to MMS have been described during research into fields as diverse as mismatch repair (Glaab *et al.*, 1998), amino acid biosynthesis (Kafer, 1987) and p53 function (Kuo *et al.*, 1997). It is also theoretically possible that a component of the putative entry pathway of MMS into *T. brucei* has been mutated resulting in decreased uptake of the mutagen and therefore an increased resistance. However, whether or not such a MMS pathway exists is unclear. The putative secondary mutation generated during the disruption of *DMC1* was not characterised here and it was judged that it was unlikely to affect further assays (see below) as any random mutation is unlikely to have occurred in a gene involved in the recombination pathways we are studying, and due to the fact that the increased resistance is only slight, with the cell lines becoming sensitive to MMS at the higher concentrations used.

4.6 Analysis of recombination efficiency in the *RAD51-3*, *RAD51-5* and *DMC1* mutants

To examine the function of the *RAD51*-like genes in *T. brucei* recombination, a transformation assay was used. This represents the only assay currently available to examine homologous recombination efficiency and demonstrated the importance of RAD51 in *T. brucei* recombination (Conway *et al.*, 2002b; McCulloch and Barry, 1999). The recombination assay involved the transformation of a bleomycin or puromycin resistance marker, flanked by tubulin intergenic sequences 240 bp and 330 bp in size, into the strains of interest (Fig. 4.21). The construct is integrated into the tubulin array by homologous recombination and, therefore, the number of transformants recovered directly relates to the efficiency of recombination in that strain. In wild type *T. brucei*, transformed DNA containing homology with the genome appears to virtually

always be recombined by HR, rather than NHEJ (Conway *et al.*, 2002a; Conway *et al.*, 2002b)

The transformation constructs were excised from the plasmid backbone by restriction digestion reactions, which were subsequently phenol:chloroform extracted, ethanol precipitated and resuspended in sterile H_2O. The transformation construct Tub-PUR-Tub was excised from pTPT (R. McCulloch, gift) by *Not*I and *Apa*I restriction digestion and the Tub-BLE-Tub construct was excised from pRM450 (R. McCulloch, gift) by *Not*I and *Apa*I digestion. Approximately 5 μg of digested DNA was used per transformation. In each transformation, 5 x 10^7 cells of each cell line were electroporated, as described in Section 2.4. The transformed cells were recovered, in all cases in non-selective media, for three generation times and 5 x 10^6 wild type and heterozygous mutant cells were plated out over 24 wells in 1.5 mls of antibiotic selective media per well. The *dmc1-/-* cell lines, where a recombination defect was not expected, were plated out as described for the wild type and heterozygous mutants. In contrast, the *rad51-3* and *rad51-5* homozygous mutants were plated out at 2 x 10^7 cells over 48 wells in 1.5 mls of antibiotic selective media per well, again allowing 3 generations for recovery and taking into account the increased doubling times (see Section 4.3.1). 3 transformations of each of the heterozygous and homozygous cell lines were carried out, along with 3 wild type control transformations. The *RAD51-3* and *RAD51-5* cell lines were transformed with Tub-BLE-Tub and selected with 2 $\mu g.ml^{-1}$ of phleomycin, whereas the *DMC1* cell lines were transformed with Tub-PUR-Tub and selected with 1 $\mu g.ml^{-1}$ of puromycin.

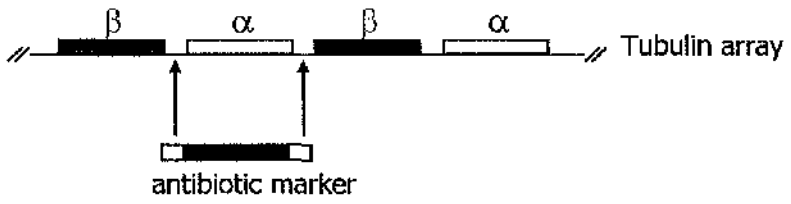


Figure 4.21: Integration of the recombination assay construct. The construct contains tubulin intergenic regions flanking a resistance marker. The construct is integrated into the tubulin array by homologous recombination, acting on the tubulin flanks, which results in an α-tubulin open reading frame being replaced with the resistance marker. Read through transcription in the tubulin array provides expression of the resistance cassette, which contains no promoter.

The number of wells containing antibiotic resistant transformants were counted after 14 days of growth and used to calculate the recombination efficiency (expressed as the

number of transformants per 10^6 cells plated out). These data are shown in Figure 4.22 and tabulated in Table 4.7.

	Cell line			
Gene	1 +/-	1 -/-	2 +/-	2 -/-
RAD51-3	1.0	3.8	1.0	9.0
RAD51-5	1.0	7.2	0.9	6.4
DMC1	0.9	1.1	1.0	0.9

Table 4.7: Recombination efficiency of the *RAD51-3*, *RAD51-5* and *DMC1* mutants. The table shows the fold reduction in the average recombination efficiency of each of the heterozygous (+/-) and homozygous (-/-) mutants when compared to the average recombination efficiency of 3174.2 wild type cells. Values of 1 represent a recombination efficiency in that cell line equivalent to wild type cells, values of less than 1 indicates a greater recombination efficiency, and values greater than 1 indicate lesser efficiency.

There is a clear reduction in recombination efficiency in the *rad51-3* and *rad51-5* homozygotes (Fig. 4.22). In the *rad51-3-/-* cells, homozygous mutant cell line 1 showed a 3.8-fold reduction relative to wild type and its cognate heterozygote, while line 2 showed a 9-fold reduction (assuming an average transformation efficiency of 4.67 in 10^6 cells for the wild type and heterozygous mutants in this experiment). Whether or not this difference is simply a statistical artefact or reflects genuine differences between the two mutants is unclear. The two *rad51-5-/-* cell lines showed comparable 6.7 and 7.7-fold reductions in transformation efficiency (assuming an average transformation efficiency of 4.38 in 10^6 cells for the wild type and heterozygous mutants in this experiment). These reductions are not as severe as that observed in *rad51-/-* mutants where a 7 to 16-fold reduction in homologous recombination efficiency is observed for a range of constructs (Conway *et al.*, 2002b). This suggests that the roles of RAD51-3 and RAD51-5 may not be as central as that of RAD51 in the process of homologous recombination, although the general efficiencies of transformation are higher in this present work. The *dmc1-/-* cell lines showed no defect in recombination compared to wild-type cells. Here, the wild type transformation efficiency is somewhat lower than that observed in the *RAD51-3* and *RAD51-5* experiments; this is probably because Tub-PUR-Tub was used rather than Tub-BLE-Tub. Either the antibiotic selection alters the level of transformant growth, or slight differences in the size of the transformed constructs subtly affect recombination efficiency.

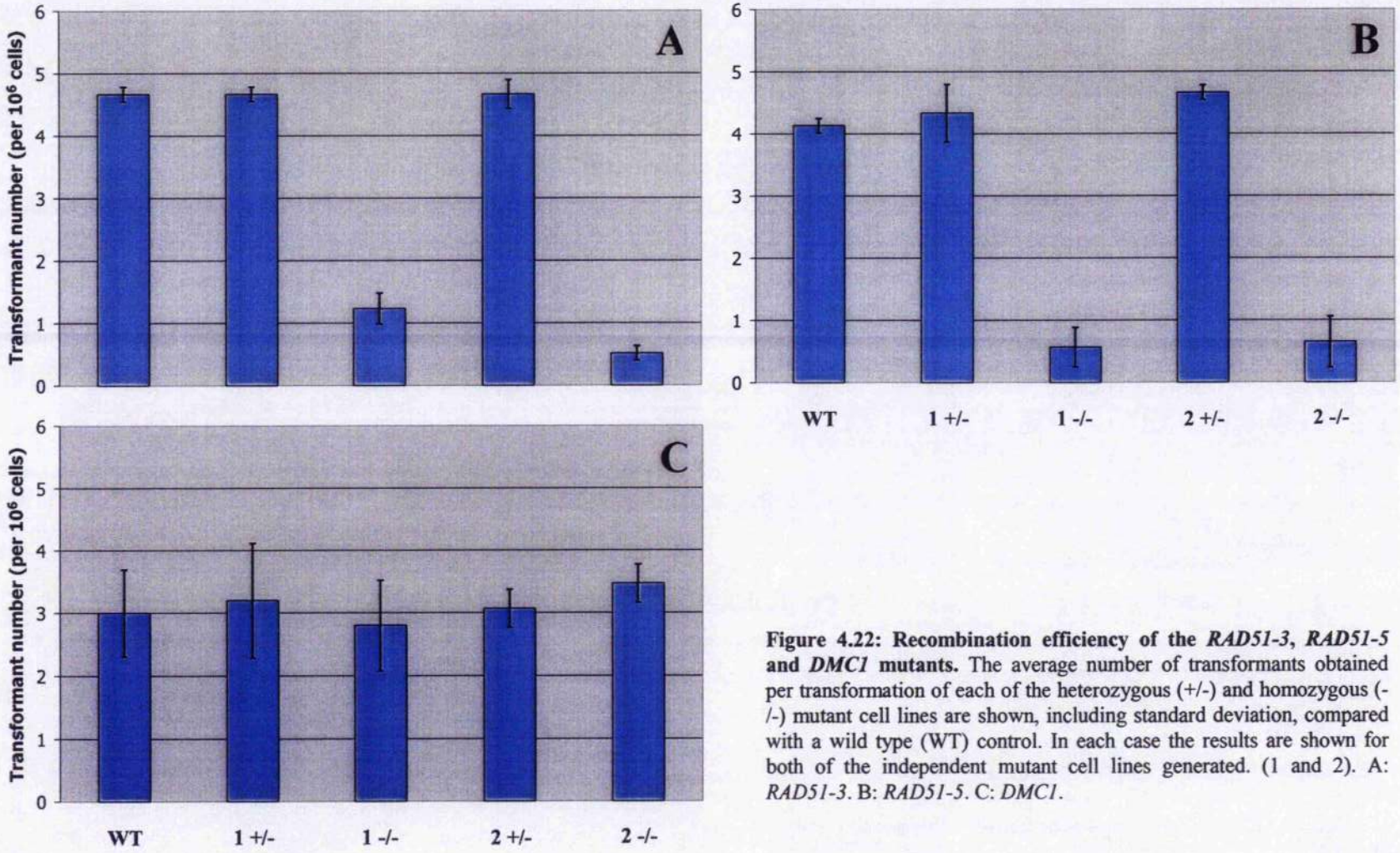


Figure 4.22: Recombination efficiency of the *RAD51-3*, *RAD51-5* and *DMC1* mutants. The average number of transformants obtained per transformation of each of the heterozygous (+/-) and homozygous (-/-) mutant cell lines are shown, including standard deviation, compared with a wild type (WT) control. In each case the results are shown for both of the independent mutant cell lines generated. (1 and 2). A: *RAD51-3*. B: *RAD51-5*. C: *DMC1*.

A number of transformants from each of the mutants were analysed by drug resistance and Southern analysis to determine the integration location of the transformed constructs. The construct can integrate by homologous recombination into three possible locations depending on the cell line being analysed. These are: the tubulin array (see Fig. 4.21), the *221VSG* ES and the mutated copy of the *RAD51*-like gene. These locations each contain the processing signals from the tubulin array identical to those contained within the constructs being used to test recombination efficiency. The same processing signals were used in the Δ*RAD51-5*::PUR first round gene disruption construct (Fig. 4.3), both *RAD51-3* constructs (Fig. 4.2) and were also used in the generation of the transgenic 3174.2 strain when the hygromycin resistance marker was integrated into the 5′ end of the *VSG221* ES (Fig. 4.28) to allow VSG switching analysis (Section 5.9). Determining the antibiotic resistance or sensitivity of the transformants can therefore assess replacement of the markers by the recombination construct.

One transformant from each wild type, heterozygous mutant and *dmc1* homozygous mutant transformations and four from the *rad51-3* and *rad51-5* homozygous transformations, were selected for further analysis. The transformants were passaged into selective media containing either hygromycin (5.0 $\mu g.ml^{-1}$), puromycin (0.25 $\mu g.ml^{-1}$) for *RAD51-3* and *RAD51-5* heterozygotes or blasticidin (2.5 $\mu g.ml^{-1}$) for the *RAD51-3* homozygotes.

The same transformant cell lines were also subjected to Southern analysis. 5 µg of genomic DNA from each transformant was restriction digested with *Hin*dIII before being separated by electrophoresis on a 0.8% agarose gel, which was then Southern blotted and probed with the bleomycin resistance open reading frame (in the case of the *RAD51-3* an *RAD51-5* transformations) or the puromycin resistance open reading frame (for the *DMC1* transformations).

Together, the results of these two assays determine the integration locus. None of the transformant cell lines had integrated the constructs by means other than homologous recombination (Figs. 4.23-25). Despite the *rad51-3-/-* and *rad51-5-/-* mutants having reduced recombination efficiency, all of the integration events analysed had occurred at one of the three available regions of homology described above. This result is slightly different to that observed in *rad51-/-* cells (Conway *et al.*, 2002b), where low levels of aberrant integrations, relying on 5-15 bp of sequence homology, take place in these assays. This may suggest that the roles of RAD51-3 and RAD51-5 in recombination are not as important as that of RAD51. Unsurprisingly, given that *dmc1-/-* mutants displayed no defect in recombination, all integration had occurred by sequence

homology. In some cases the transformants displayed two bands in the Southerns: this is as a result of two transformants being selected for in the same well, resulting in a polyclonal population. Genomic DNA was prepared from transformants, after selection using the transformed resistance cassette, in the absence of further drug selection, meaning that there was no selection against polyclonal populations. In all cases, the polyclonal populations occurred in wild type cells or mutant cell lines without a recombination deficiency, suggesting that in those cell lines the recombination efficiency may actually be higher than calculated previously (Fig. 4.22).

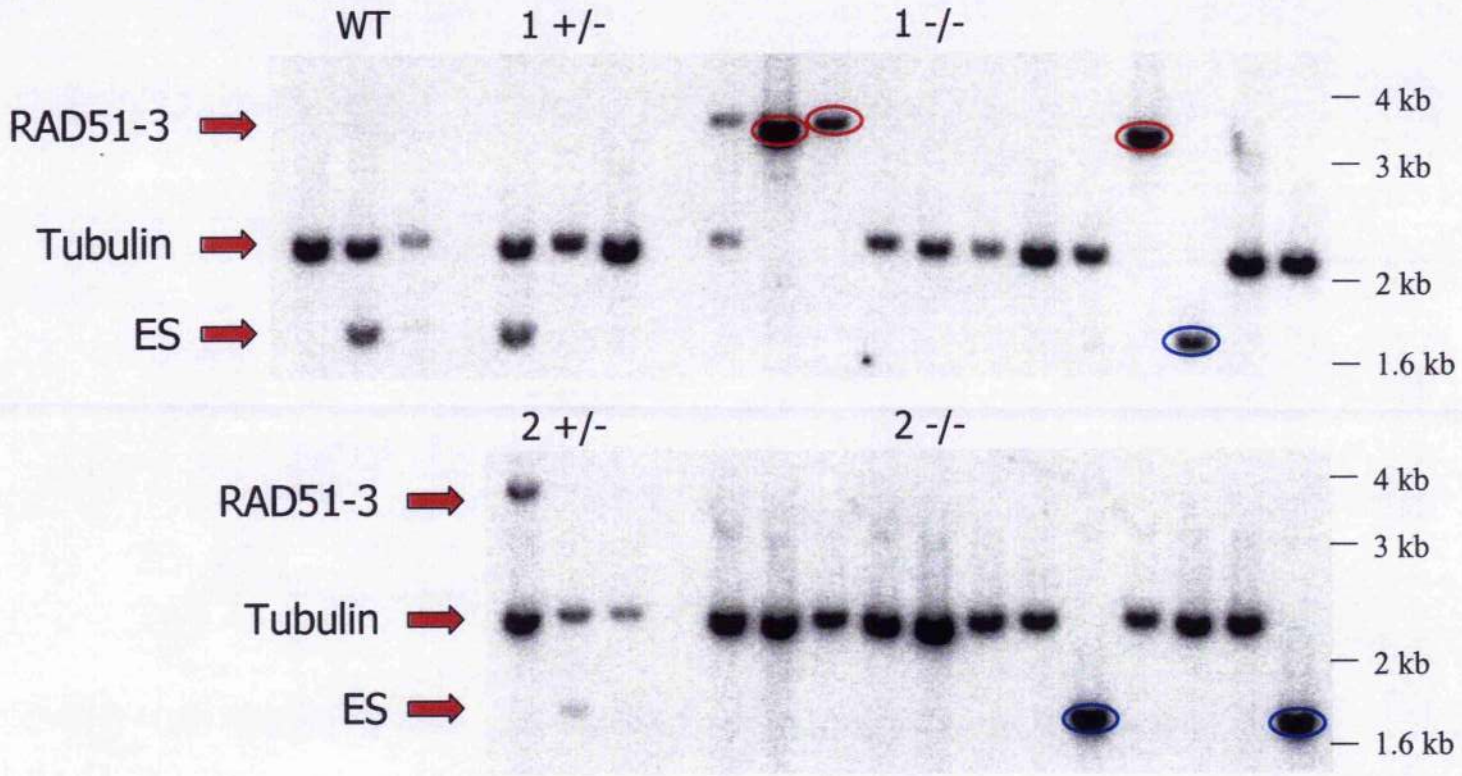


Figure 4.23: Analysis of construct integration in the *RAD51-3* mutants. Two Southern blots are shown of Tub-BLE-Tub transformants digested with *Hin*dIII and probed with the BLE open reading frame. A number of transformants are shown for wild type (WT) cells, 2 independent *RAD51-3* heterozygous mutants (+/-) and *RAD51-3* homozygous mutants (-/-). Bands representing the expected fragment size of Tub-BLE-Tub integrated into the tubulin array (Tubulin), the 221VSG expression site (ES) and the disrupted RAD51-3 gene (RAD51-3) are indicated. Transformants with two bands represent polyclonal populations. The bands circled blue correspond to transformants that are sensitive to hygromycin while those in red are sensitive to blasticidin or puromycin.

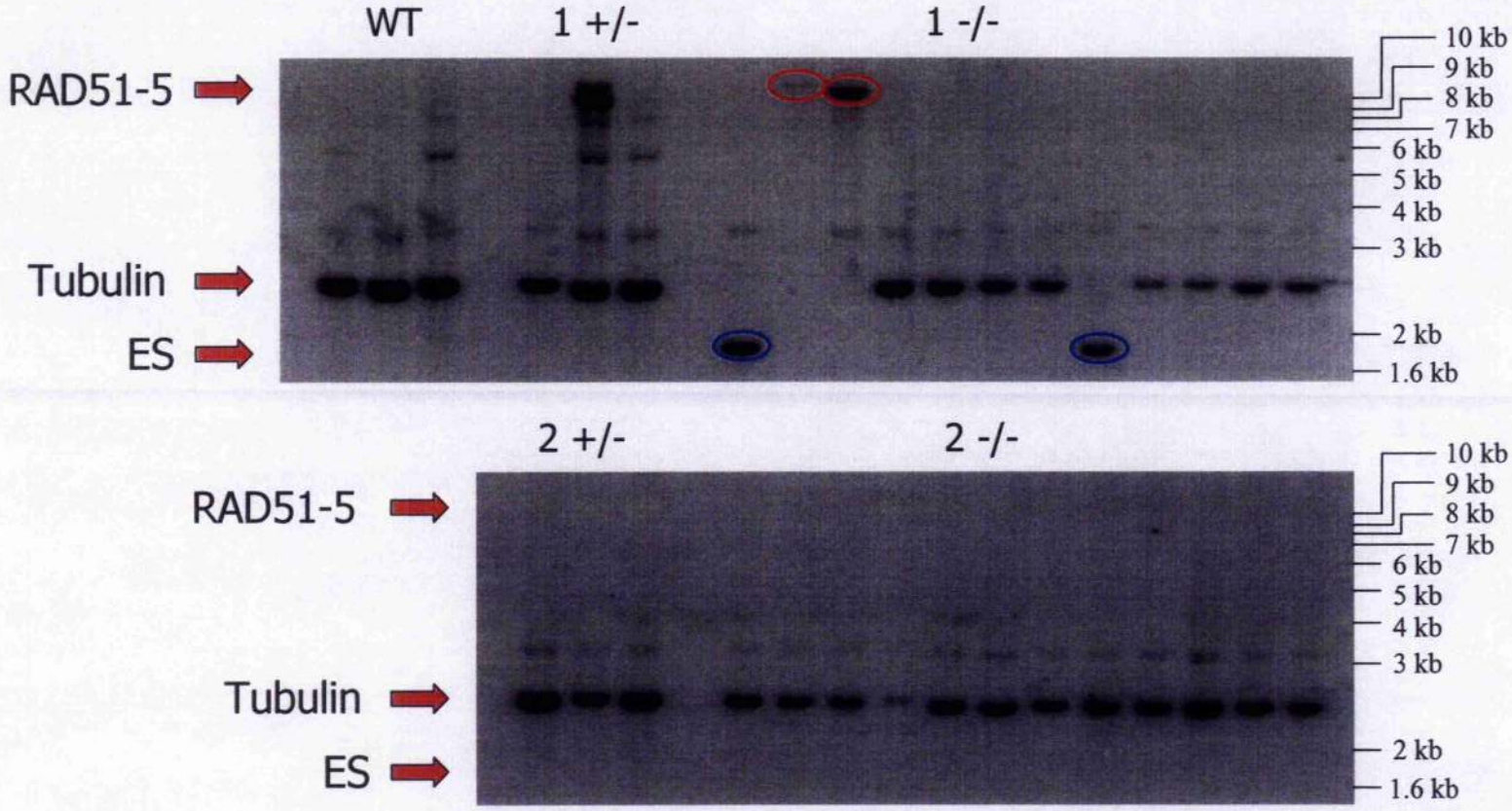


Figure 4.24: Analysis of construct integration in the *RAD51-5* mutants. Two Southern blots are shown of Tub-BLE-Tub transformants digested with *Hin*dIII and probed with the BLE open reading frame. A number of transformants are shown for wild type (WT) cells, 2 independent *RAD51-5* heterozygous mutants (+/-) and *RAD51-5* homozygous mutants (-/-). Bands representing the expected fragment size of Tub-BLE-Tub integrated into the tubulin array (Tubulin), the 221VSG expression site (ES) and the disrupted RAD51-5 gene (RAD51-5) are indicated. Transformants with two bands represent polyclonal populations. The bands circled blue correspond to transformants that are sensitive to hygromycin while those in red are sensitive to blasticidin or puromycin.

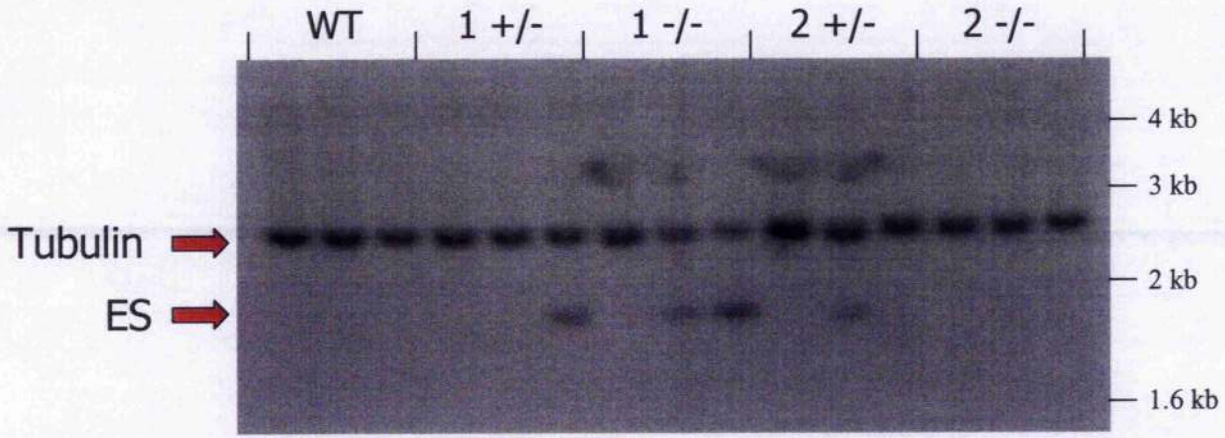


Figure 4.25: Analysis of construct integration in the *DMC1* mutants. Two Southern blots are shown of Tub-PUR-Tub transformants digested with *Hin*dIII and probed with the PUR open reading frame. A number of transformants are shown for wild type (WT) cells, 2 independent *DMC1* heterozygous mutants (+/-) and *DMC1* homozygous mutants (-/-). Bands representing the expected fragment size of Tub-BLE-Tub integrated into the tubulin array (Tubulin) and the 221VSG expression site (ES) are indicated. Transformants with two bands represent polyclonal populations.

4.7 Generation of re-expression cell lines

In order to examine the effect of the *RAD51*-like genes on VSG switching, the generation of re-expression cell lines was necessary to prove that any defect observed was purely as a result of loss of the functional gene and not a secondary mutation. One of the two independent homozygous mutants of *RAD51-3* and *RAD51-5* were used to generate a re-expression cell line. A re-expression cell line for *DMC1* was not generated, as loss of this gene had no defect in any of the assays carried out thus far and we deemed it unlikely that it would not have a role in VSG switching.

4.7.1 Generation of *RAD51-3* and *RAD51-5* re-expression cell lines

The *RAD51-3* and *RAD51-5* genes were PCR-amplified from 3174.2 genomic DNA (primers *RAD51-3*-For and *RAD51-3*-Rev, or *RAD51-5*-For and *RAD51-5*-Rev) using Herculase DNA polymerase, restriction digested (*RAD51-3* was digested with *Eco*RV, and *RAD51-5* was digested with *Eco*RI) and made blunt using the Klenow fragment of *E. coli* DNA polymerase before being ligated into the plasmid pRM481 (R. McCulloch, gift), which had been *Eco*RV-digested and CIP treated (Fig 4.26). This resulted in the generation of pRM481::*RAD51-3* and pRM481::*RAD51-5*. Insertion of *RAD51-3* or *RAD51-5* into pRM481 allows the genes to be inserted into the tubulin array, where they are transcribed from the endogenous transcription. Splicing and polyadenylation is provided by the 5′ actin and 3′ tubulin intergenic sequences, rather than the natural processing of the genes.

The constructs were excised from the plasmid backbone by restriction digestion with *Not*I and *Apa*I and the resulting digested DNA phenol:chloroform extracted and ethanol precipitated. Approximately 5 μg of digested DNA (quantified by gel electrophoresis relative to life technologies 1 kb size ladder) was then used in each *T. brucei* transformation. Two transformations of 5 x 10^7 cells were carried out on *rad51-3-/-* 2 with pRM481::*RAD51-3* and *rad51-5-/-* 1 with pRM481::*RAD51-5* using the protocol described in section 2.4. Antibiotic resistant transformants were selected for by plating out 2 x 10^7 *rad51-3-/-* cells and 3 x 10^7 *rad51-5-/-* cells from each transformation at 2 $\mu g.ml^{-1}$ phleomycin, and 1 x 10^7 cells were plated out over 12 wells with 1.5 mls per well. 5 putative *RAD51-3* re-expressers and 5 putative *RAD51-5* re-expressers were generated.

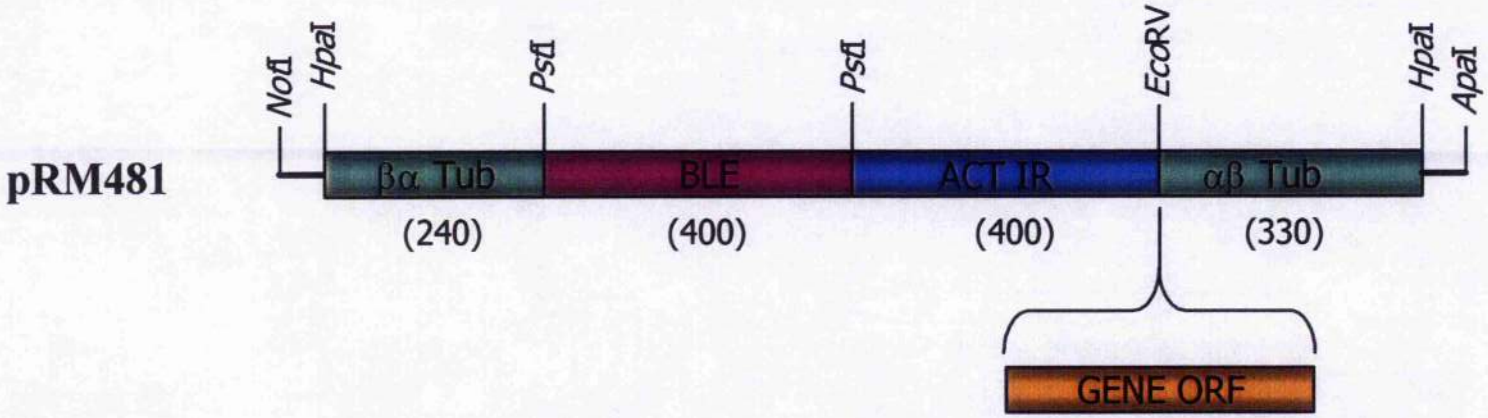


Figure 4.26: Re-Expression constructs for *RAD51-3* and *RAD51-5*. The genes are cloned into an *Eco*RV site between the actin intergenic (ACT IR) and αβ Tubulin (αβ Tub) intergenic sequences of plasmid pRM481. The construct is flanked with tubulin intergenic regions (βα Tub and αβ Tub) which allow homologous integration into the tubulin array, replacing an α tubulin open reading frame. The sizes of the constituent components are shown in base pairs, except for the inserted *RAD51-3* or *RAD51-5* genes (orange box which is not drawn to scale).

The generation of re-expresser cell lines was confirmed by PCR-amplification of the entire open reading frame using Taq DNA polymerase and primers Exp-For and Exp-Rev (data not shown; see Section 4.7.2) and one was chosen for each gene (named *RAD51-3-/-/+* and *RAD51-5-/-/+*).

4.7.2 Confirmation of the generation of *RAD51-3* and *RAD51-5* re-expression cell lines

To confirm that pRM481::*RAD51-3* and pRM481::*RAD51-5* had integrated as expected in the *RAD51-3-/-/+* and *RAD51-5-/-/+* cell lines, Southern analysis was conducted in the same manner as that used during the generation of the respective mutants (Section 4.4.5). Genomic DNA from wild type cells and the progenitor heterozygous and homozygous cell lines used to generate the re-expression line were also subjected to Southern analysis to allow direct comparison (Fig. 4.27). The extra fragments observed in the re-expression cell lines correspond to expected integration of pRM481::*RAD51-3* and pRM481::*RAD51-5* into the tubulin array, since they are of the sizes expected: *RAD51-3*, 3410bp; and *RAD51-5*, 7910bp.

4.7.3 *in vitro* growth of the *RAD51-3* and *RAD51-5* re-expression cell lines

Analysis of *in-vitro* growth of the *RAD51-3-/-/+* and *RAD51-5-/-/+* cells was carried out to determine if transcription of the genes alleviated the growth defect that was observed in the homozygous mutants (Section 4.3.1). The assay was carried out in the same manner as described in Section 4.3.1. Three repetitions of the growth assay were carried out for each re-expresser, and two repetitions of wild type cells for comparison (Fig. 4.27). In both cases, the re-expresser cell lines had population doubling times essentially equivalent to that of wild-type cells. The wild types doubled in 7.7 hrs, *RAD51-3-/-/+* in 8.2 hrs and *RAD51-5-/-/+* in 7.9 hrs.

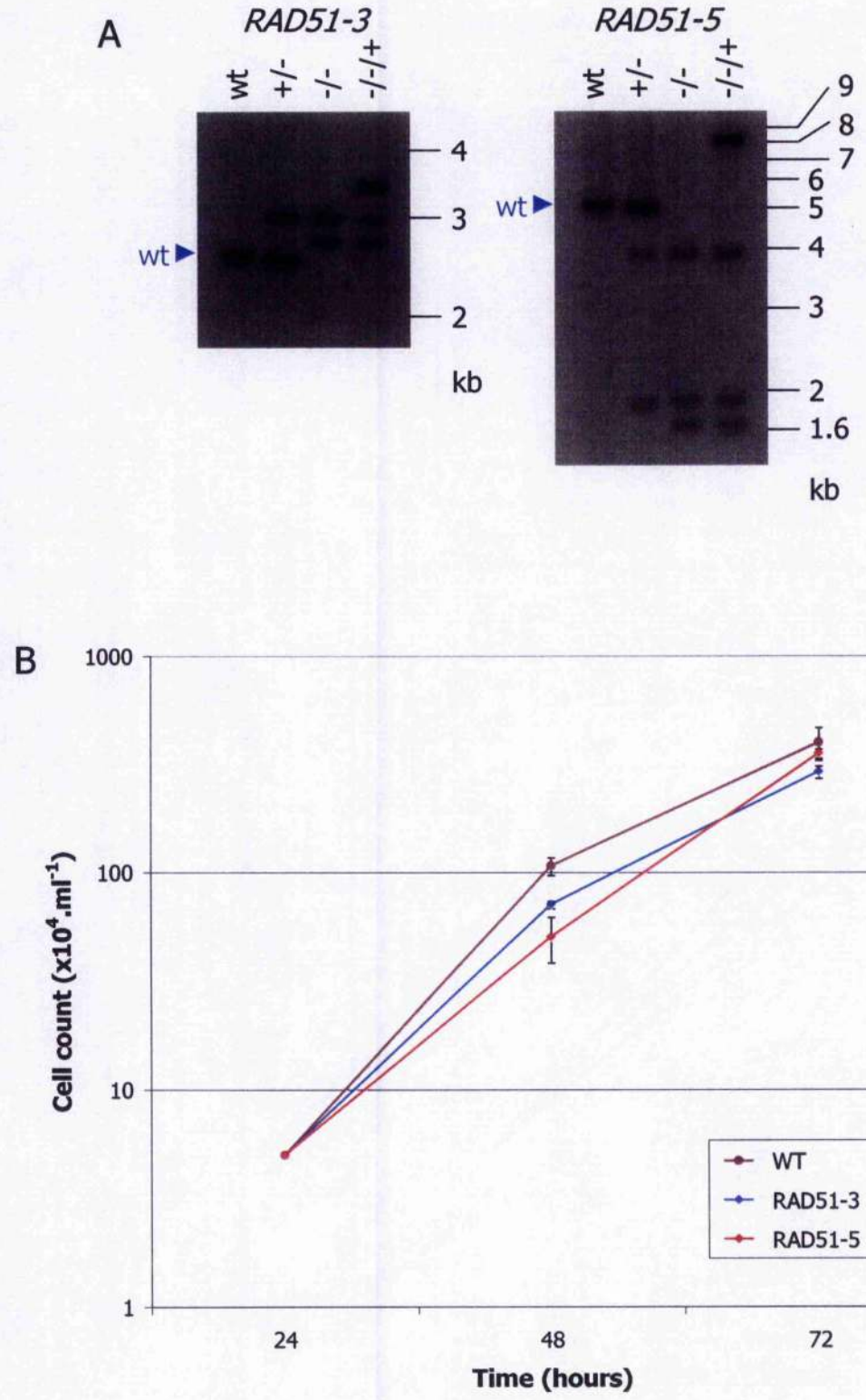


Figure 4.27: Generation of the *RAD51-3* and *RAD51-5* re-expression cell lines. A: Southern blots of 3174.2 wild type (WT) cell and either *RAD51-3* or *RAD51-5* heterozygous (+/-), homozygous (-/-) and re-expression (-/-/+) genomic DNA probed with the respective open reading frame. For *RAD51-3* the DNA was digested with *Bst*XI and for *RAD51-5* the DNA was digested with *Spe*I. The size of the fragment that corresponds to the undisrupted gene is shown (wt), as are size markers. **B:** Cell counts of 3174.2 wild type cells (WT) compared with *RAD51-3* and *RAD51-5* re-expressers (-/-/+) over 72 hours. Standard deviations are shown.

4.8 Analysis of *VSG* switching in the *RAD51-3*, *RAD51-5* and *DMC1* mutants

VSG switching assays were carried out to determine if the disruption of any of the *RAD51*-like genes had any effect on antigenic variation. The ability to carry out this analysis is as a result of the generation of the mutant cell lines in the transgenic 3174.2 strain. This trypanosome strain contains a modified active expression site, into which antibiotic resistance markers for hygromycin and G418 have been inserted (Fig 4.28), which allows its use in experiments to determine the VSG switching frequency and the relative contributions of gene conversion and *in situ* switching mechanisms. The exact conditions used in this assay are described in Sections 2.4 and 4.8.1-2, but a general description is given below. Growth on both hygromycin and G418 ensures that all cells express the modified expression site and hence the VSG221 protein. These cells were used to immunise mice against VSG221.

The *RAD51-3*, *RAD51-5* and *DMC1* cell lines, previously grown on hygromycin and G418, were removed from antibiotic selection for 9 generation times, thereby allowing VSG switching to take place, before injection into immunised mice. 24 hrs after injection, surviving trypanosomes, those that had not been killed by immune lysis as a result of switching their VSG coat, were recovered from the mice and plated out over 96 wells. After a maximum of 4 weeks the number of wells that showed trypanosome growth were counted and the VSG frequency estimated (Section 4.8.1). In addition to this, cells from a number of wells showing growth had their drug sensitivities assessed and genomic DNA was isolated from the switched trypanosomes for PCR analyses. The results of the PCR-amplification of the *VSG221* gene and the antibiotic markers, together with the switched variants sensitivity or resistance to hygromycin and G418, allowed us to determine the switching mechanism utilised in that cell line (Section 4.8.2).

4.8.1 Analysis of *VSG* switching frequency in the *RAD51-3*, *RAD51-5* and *DMC1* mutants

When analysing the VSG switching of the wild type 3174.2 cells, all the heterozygous mutants and the *dmc1-/-* mutants, 4 x 10^7 cells were injected into the immune mice. For the *rad51-3-/-* and *rad51-5-/-* mutants, where a switching defect was expected, 8 x 10^7 cells were injected in an attempt to increase the number of switching events per mouse,

making the assay more accurate. In each cell line this was achieved by inoculating 25 ml of media (2 x 25ml for the *rad51-3-/-* and *rad51-5-/-* mutants) with 1 x 10^5 cells which, after 9 generation times, reached a density of approximately 2 x 10^6 cell.ml^{-1}. By ensuring that all the wild type, heterozygous and homozygous mutants have the same time scale to allow VSG switch variants to arise, variability in the assay is minimised as much as possible. All individual switching experiments used independently grown 25 ml populations of cells allowed to undergo VSG switching.

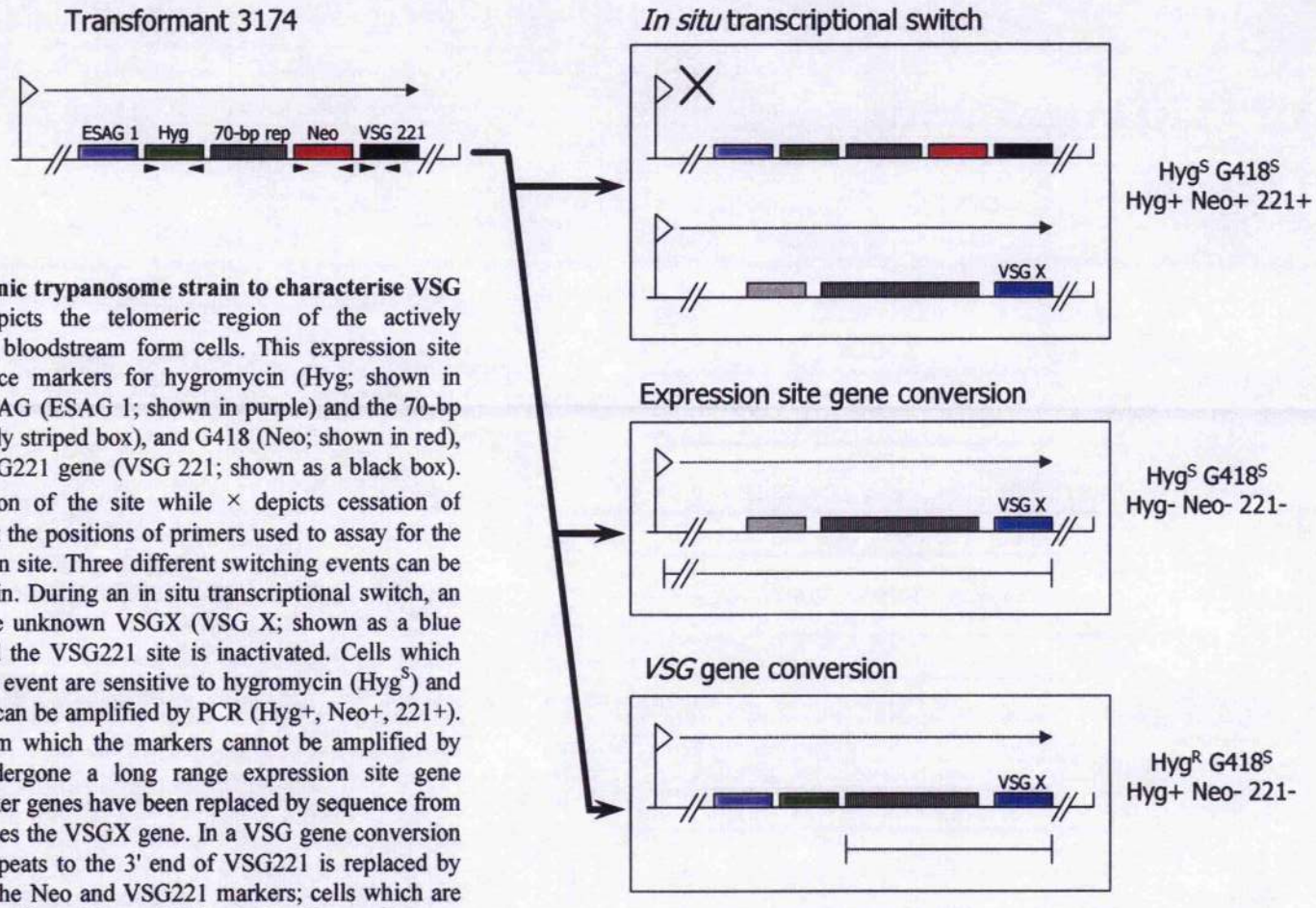


Figure 4.28: Use of the 3174 transgenic trypanosome strain to characterise VSG switching events. This scheme depicts the telomeric region of the actively transcribed expression site in 3174.2 bloodstream form cells. This expression site was modified to include the resistance markers for hygromycin (Hyg; shown in green), between the most proximal ESAG (ESAG 1; shown in purple) and the 70-bp repeats (70-bp rep; shown as a vertically striped box), and G418 (Neo; shown in red), between the 70-bp repeats and the VSG221 gene (VSG 221; shown as a black box). A dashed arrow indicates transcription of the site while × depicts cessation of transcription. Small triangles represent the positions of primers used to assay for the presence of genes within the expression site. Three different switching events can be distinguished in this trypanosome strain. During an in situ transcriptional switch, an inactive expression site containing the unknown VSGX (VSG X; shown as a blue box) is transcriptionally activated and the VSG221 site is inactivated. Cells which have undergone this type of switching event are sensitive to hygromycin (Hyg^S) and G418 ($G418^S$), and each marker gene can be amplified by PCR (Hyg+, Neo+, 221+). Cells which are Hyg^S and $G418^S$ from which the markers cannot be amplified by PCR (Hyg-, Neo-, 221-), have undergone a long range expression site gene conversion event in which all the marker genes have been replaced by sequence from another expression site which introduces the VSGX gene. In a VSG gene conversion event the sequence from the 70-bp repeats to the 3' end of VSG221 is replaced by sequence containing VSGX, deleting the Neo and VSG221 markers; cells which are resistant to hygromycin (Hyg^R) but $G418^S$, and from which only the Hyg marker can be amplified by PCR (Hyg+, Neo-, 221-) have undergone this type of switching event. The integrity of the genomic DNA template was confirmed using primers directed against the large subunit of RNA polymerase I (Rudenko et al., 1996). (After McCulloch and Barry, 1999).

The VSG switching frequency was calculated for each cell line in each mouse as follows: the number of wells showing growth for each mouse was first multiplied by 2.5 to calculate the number of switched cells in the total blood volume (the total blood volume of a mouse is taken as 2 ml, and 2 x 0.4 ml of blood was used to isolate switched variants). This number was then divided by the of number of generation times that had occurred during the 24 hour period following infection in the immunised mouse, thus determining the number of switched cells in the injected sample. Finally, the number of cells that were injected was taken into account and the number of switched variants per 10^7 cells injected was determined. These results are shown in Tables 4.8, 4.10 and 4.12. Statistical analysis was also carried out on the results to determine if any observed differences were significant or not (Tables 4.9, 4.11 and 4.13). It should be noted that, in calculating the number of generations the switched variants underwent in each mouse, the *in vitro* population doubling times were used (refer to Table 4.5, Section 4.3.1), meaning that the longer generation times of the *rad51-3* and *rad51-5* homozygous mutants is taken into account. As has been argued, these data were used in preference to the *in vivo* population doubling times (Table 4.6, section 4.3.2), as they are more accurate.

From these results, it appears that only one of the *RAD51*-like genes is involved in VSG switching. Despite the apparently equivalent importance of RAD51-3 and RAD51-5 in previous DNA repair and recombination assays, the *rad51-3-/-* mutants showed a reduced VSG switching frequency relative to wild type or heterozygous mutants, whereas the *rad51-5-/-* mutants did not (Fig. 4.29). The observed switching frequencies in the wild type and heterozygous mutants were comparable with that described previously (Bell and McCulloch, 2003; McCulloch and Barry, 1999; Robinson *et al.*, 2002): around 5 switched variants in 10^7 cells, although there was considerable variation. The *rad51-5-/-* mutants also showed this rate and statistical analysis confirmed there was no statistical difference ($p>0.05$). In contrast, the *rad51-3-/-* mutants switched at a lower rate: this was, on average, 0.5-1.5 switched variants per 10^7 cells, but varied from 0.1-2.6. In fact, in a number of experiments, no switched variants were recovered, suggesting that the frequency may be at the threshold of detection with the number of cells being injected.

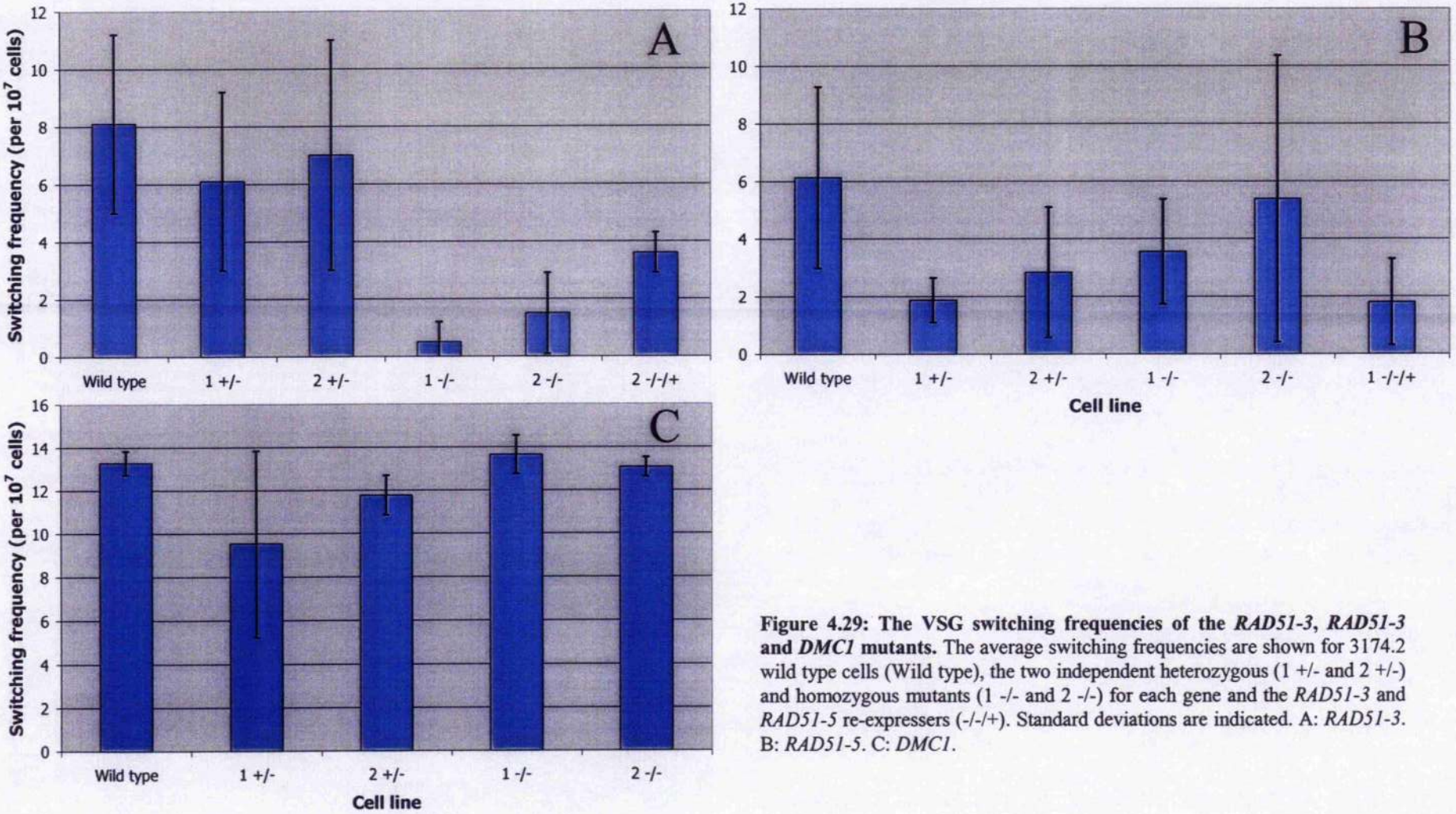


Figure 4.29: The VSG switching frequencies of the *RAD51-3*, *RAD51-3* and *DMC1* mutants. The average switching frequencies are shown for 3174.2 wild type cells (Wild type), the two independent heterozygous (1 +/- and 2 +/-) and homozygous mutants (1 -/- and 2 -/-) for each gene and the *RAD51-3* and *RAD51-5* re-expressers (-/-/+). Standard deviations are indicated. A: *RAD51-3*. B: *RAD51-5*. C: *DMC1*.

These negative experiments were not, however, included in the frequency calculation or statistical examination, which suggest that the reduction in VSG switching is significant ($p<0.05$). Comparing the switch frequency of either *rad51-3-/-* mutant lines to that of the wild type cells was on the verge of significance (p values of 0.05 and 0.07), whereas comparison with the heterozygous mutants from which they were derived, and to the re-expresser, always suggested a significant effect.

Unsurprisingly, given the lack of an observable function for *DMC1* in DNA repair or recombination in 3174.2 bloodstream form cells, no effect was observed on VSG switching frequency by mutating the gene. For reasons that are unclear, the observed switch frequencies in this experimental set were nearly uniformly higher (around 10 switched variants in 10^7 cells) than the RAD51-3 or RAD51-5 experiments, but disruption of *DMC1* had no effect.

RAD51-3	1 +/-	1 -/-	2 +/-	2 -/-	2 -/-/+
Wild type	0.442	0.052	0.661	0.073	0.133
1 +/-		0.008	0.702	0.014	0.117
1 -/-			0.012	0.258	0.003
2 +/-				0.021	0.105
2 -/-					0.046

Table 4.9: Statistical analysis of the VSG switching frequencies in the *RAD51-3* mutants. P values are shown for two sample T-tests comparing the VSG switching frequencies of 3174.2 wild type cells (wild type), the two *RAD51-3* heterozygous mutants (1 +/- and 2 +/-), the two homozygous mutants (1 -/- and 2 -/-) and the *RAD51-3* re-expresser (-/-/+).

Cell line	No. of wells (out of 192)	Switch frequency (per 10^7 cells)	Average	St Dev
Wild type	149	11.6		
	84	6.6		
	D			
	78	6.1	**8.1**	3.1
1 +/-	93	7.3		
	127	9.9		
	84	6.6		
	63	4.9		
	9	0.7		
	96	7.5	**6.1**	3.1
2 +/-	152	11.9		
	139	10.9		
	104	8.1		
	31	2.4		
	76	5.9		
	33	2.6	**7.0**	4.0
1 -/-	4	0.4		
	2	0.2		
	0			
	0			
	0			
	1	0.1		
	16	1.7		
	1	0.1	**0.5**	0.7
2 -/-	D			
	27	2.8		
	0			
	0			
	0			
	2	0.2		
	4	0.4		
	25	2.6	**1.5**	1.4
2 -/-/+	56	4.4		
	40	3.1		
	43	3.4	**3.6**	0.7

Table 4.8: The VSG switching frequencies of the *RAD51-3* mutants. The numbers of wells showing growth are shown for each of the switching assays carried out on wild type 3174.2 cells, each independent *RAD51-3* heterozygous (1 +/- and 2 +/-) and homozygous mutant (1 -/- and 2 -/-) and the *RAD51-3* re-expression cell line. The table also includes the calculated switching frequency for each experiment, the average switching frequency for each cell line and the standard deviation (St Dev). D: death of the mouse during the experiment. 0: no wells showed growth.

Cell line	No. of wells (out of 192)	Switch frequency	Average	St Dev
Wild type	107	8.4		
	50	3.9	**6.1**	3.1
1 +/-	D			
	31	2.4		
	17	1.3	**1.9**	0.8
2 +/-	26	2.0		
	69	5.4		
	14	1.1	**2.8**	2.3
1 -/-	0			
	62	4.8		
	29	2.3	**3.6**	1.8
2 -/-	0			
	114	8.9		
	24	1.9	**5.4**	5.0
1 -/-/+	10	0.8		
	45	3.5		
	14	1.1	**1.8**	1.5

Table 4.10: The VSG switching frequencies of the *RAD51-5* mutants. The numbers of wells showing growth are shown for each of the switching assays carried out on wild type 3174.2 cells, each independent *RAD51-5* heterozygous (1 +/- and 2 +/-) and homozygous mutant (1 -/- and 2 -/-) and the *RAD51-5* re-expression cell line. The table also includes the calculated switching frequency for each experiment, the average switching frequency for each cell line and the standard deviation (St Dev). D: death of the mouse during the experiment. 0: no wells showed growth.

RAD51-5	1 +/-	1 -/-	2 +/-	2 -/-	1 -/-/+
Wild type	0.315	0.497	0.424	0.886	0.322
1 +/-		0.431	0.560	0.499	0.956
1 -/-			0.730	0.706	0.454
2 +/-				0.617	0.566
2 -/-					0.500

Table 4.11: Statistical analysis of the VSG switching frequencies in the *RAD51-5* mutants. P values are shown for two sample T-tests comparing the VSG switching frequencies of 3174.2 wild type cells (wild type), the two *RAD51-5* heterozygous mutants (1 +/- and 2 +/-), the two homozygous mutants (1 -/- and 2 -/-) and the *RAD51-5* re-expresser (-/-/+).

Cell line	No. of wells (out of 192)	Switch frequency	Average	St Dev
Wild type	165	12.9		
	175	13.7	**13.3**	0.6
1 +/-	D			
	83	6.5		
	161	12.6	**9.5**	4.3
2 +/-	146	11.4		
	164	12.8		
	142	11.1	**11.8**	0.9
1 -/-	184	14.4		
	178	13.9		
	162	12.7	**13.6**	0.9
2 -/-	174	13.6		
	164	12.8		
	164	12.8	**13.1**	0.5

Table 4.12: The VSG switching frequencies of the *DMC1* mutants. The numbers of wells showing growth are shown for each of the switching assays carried out on wild type 3174.2 cells and each independent *DMC1* heterozygous (1 +/- and 2 +/-) and homozygous mutant (1 -/- and 2 -/-). The table also includes the calculated switching frequency for each experiment, the average switching frequency for each cell line and the standard deviation (St Dev). D: death of the mouse during the experiment.

DMC1	1 +/-	1 -/-	2 +/-	2 -/-
Wild type	0.437	0.626	0.146	0.712
1 +/-		0.410	0.604	0.456
1 -/-			0.080	0.370
2 +/-				0.158

Table 4.13: Statistical analysis of the VSG switching frequencies in the *DMC1* mutants. P values are shown for two sample T-tests comparing the VSG switching frequencies of 3174.2 wild type cells (wild type), the two *DMC1* heterozygous mutants (1 +/- and 2 +/-) and the two homozygous mutants (1 -/- and 2 -/-).

4.8.2 Analysis of *VSG* switching mechanism in the *RAD51-3*, *RAD51-5* and *DMC1* mutants

The 3174.2 cell line allows us to distinguish three different types of switching event that results in the expression of a novel VSG: *in situ* switching, VSG gene conversion and expression site gene conversion (Fig. 4.28). Cell lines that have carried out an *in situ* switch are sensitive to both hygromycin (Hyg^S) and G418 ($G418^S$) as the resistance genes are no longer expressed, but PCR products are still obtained for each of the markers (Hyg+, Neo+, 221+) as no loss of sequence occurred during the switch. Cell lines that have carried out an expression site gene conversion are also both Hyg^S and $G418^S$, but here PCR products are not obtained for any of the markers (Hyg-, Neo-, 221-) as they have been removed and replaced by sequence from another expression site. The third type of switching event, *VSG* gene conversion, results in cells that are hygromycin resistant (Hyg^R) but $G418^S$. *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; this removes the *Neo* and *VSG221* markers, but leaves the Hyg intact (Hyg+, Neo-, VSG21-).

Where possible, at least 10 switched variants from each cell line and from each mouse were analysed to determine their switching mechanism (refer to Sections 2.4.2). The antibiotic resistance or sensitivity of each switched variant was scored using hygromycin (5.0 $\mu g.ml^{-1}$) or neomycin (2.5 $\mu g.ml^{-1}$). PCR-amplification of the hygromycin (primers Hygro-5′ and Hygro-3′; Appendix 1), neomycin (primers Neo-5′ and Neo-3′; Appendix 1) and VSG221 (primers VSG221-5′ and VSG221-3′; Appendix 1) open reading frames within the modified expression site was used to determine the presence, or absence, of each of the genes in the *VSG221* expression site. RNA polymerase I (primers Pol-5′ and Pol-3′; Appendix 1) reactions were carried out to control for the presence of intact genomic DNA. Together, the results for the antibiotic selection and PCR were used to determine the switching mechanisms used in each *RAD51*-like mutant and are shown in Figures 4.30-32. It appears that disruption of none of the 3 *RAD51*-like genes had a major effect on the type of switching event that can be used. There is substantial variability in this work but, in general, in the wild type cells *in situ* switching and expression site gene conversion occurred at relatively similar frequencies and *VSG* gene conversion, to replace just the *VSG221* gene and the upstream neomycin marker, were relatively rare. Very small numbers of events that do not conform to these events were also detected, but these are very rare.

Given this variability, it is very difficult to determine if any slight shift in the relative use of the switching events resulted from the disruption of the *RAD51*-like genes. It is more appropriate to attempt to define a major shift in switching mechanism, for example if the cell line showed a total loss of ability to carry out gene conversion switches. With this in mind, all of the mutations appear to have no effect on the VSG switch mechanism. For the *rad51-5* and *dmc1* homozygous mutants, the lack of altered switching frequency is reflected in a lack of alteration in switch mechanism. For *rad51-3-/-* mutants, the reduced VSG switch frequency is not reflected in an altered pattern of switch mechanism. This is most readily explained by an equal reduction in all the VSG switching mechanisms following the disruption of *RAD51-3*.

	In situ	ES GC	VSG GC	Other	Total
Wild type	9	1	0	0	10
Total	**9**	**1**	**0**	**0**	**10**
	2	8	0	0	10
1 +/-	4	2	2	1	9
	1	9	0	0	10
Total	**7**	**19**	**2**	**1**	**29**
	0	1	0	0	1
1 -/-	3	11	2	0	16
	0	1	0	0	1
Total	**3**	**13**	**2**	**0**	**18**
	6	1	3	0	10
2 +/-	7	3	0	0	10
	1	6	3	0	10
Total	**14**	**10**	**6**	**0**	**30**
	0	1	0	0	1
2 -/-	3	0	0	0	3
	20	3	0	0	23
Total	**23**	**4**	**0**	**0**	**27**
	10	0	0	0	10
2 -/-/+	4	5	0	1	10
	7	1	1	1	10
Total	**21**	**6**	**1**	**2**	**30**

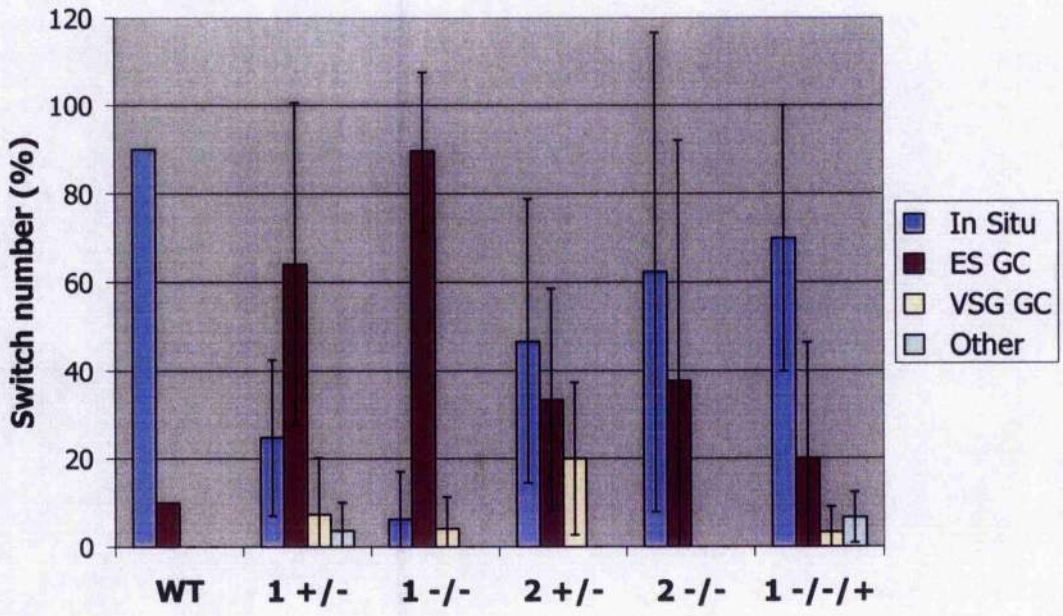


Figure 4.30: Analysis of switching events in the *RAD51-3* mutants. In the table (top) the type of switching mechanism used in a number of putatively clonal switched variants, from each of the *RAD51-3* cell lines and 3174.2 wild type cells, is indicated. The data are shown for the individual experiments detailed in the text. The graph (bottom) shows the tabulated data averaged for each cell line, and standard deviations are indicated (except for the wild type cells as switched variants from only one mouse were analysed). For both parts, the following nomenclature is used: In situ: in situ transcriptional switch. ES GC: expression site gene conversion. VSG GC: VSG gene conversion. Other: unknown. WT: wild type. +/-: heterozygous mutant. -/-: homozygous mutant. /-/-+: re-expression cell line.

	In situ	ES GC	VSG GC	Other	Total
Wild type	3	6	0	1	10
	2	4	3	1	10
Total	**5**	**10**	**3**	**2**	**20**
1 +/-	6	4	0	0	10
	6	4	0	0	10
Total	**12**	**8**	**0**	**0**	**20**
1 -/-	7	3	0	0	10
	10	0	0	0	10
Total	**17**	**3**	**0**	**0**	**20**
	0	6	3	1	10
2 +/-	3	6	0	1	10
	7	3	0	0	10
Total	**10**	**15**	**3**	**2**	**30**
2 -/-	6	4	0	0	10
	8	1	1	0	10
Total	**14**	**5**	**1**	**0**	**20**
1 -/-/+	9	0	0	1	10
	7	2	1	0	10
Total	**16**	**2**	**1**	**1**	**20**

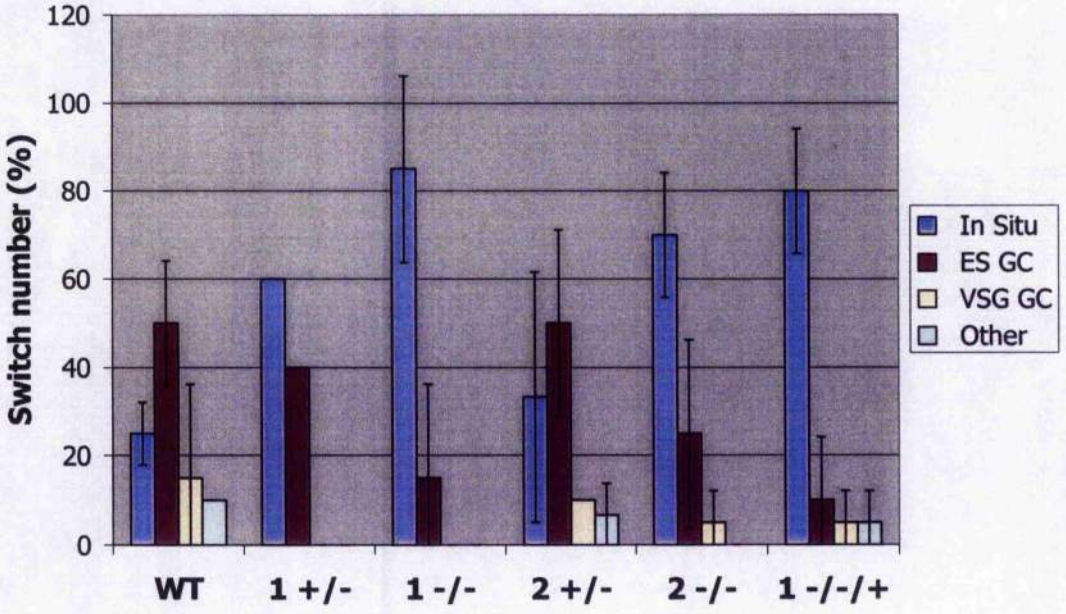


Figure 4.31: Analysis of switching events in the *RAD51-5* mutants. In the table (top) the type of switching mechanism used in a number of putatively clonal switched variants, from each of the *RAD51-5* cell lines and 3174.2 wild type cells, is indicated. The data are shown for the individual experiments detailed in the text. The graph (bottom) shows the tabulated data averaged for each cell line, and standard deviations are indicated. For both parts, the following nomenclature is used: In situ: in situ transcriptional switch. ES GC: expression site gene conversion. VSG GC: VSG gene conversion. Other: unknown. WT: wild type. +/-: heterozygous mutant. -/-: homozygous mutant. /-/-+: re-expression cell line.

	In-situ	ES GC	VSG GC	Other	Total
Wild-type	5	3	2	0	10
	1	6	3	0	10
Total	**6**	**9**	**5**	**0**	**20**
1 +/-	3	5	2	0	10
	5	4	1	0	10
Total	**8**	**9**	**3**	**0**	**20**
	6	4	0	0	10
1 -/-	5	1	4	0	10
	5	2	2	1	10
Total	**16**	**7**	**6**	**1**	**30**
	5	1	4	0	10
2 +/-	5	3	2	0	10
	5	2	2	0	9
Total	**15**	**6**	**8**	**0**	**29**
	7	3	0	0	10
2 -/-	6	2	1	1	10
	8	2	0	0	10
Total	**21**	**7**	**1**	**1**	**30**

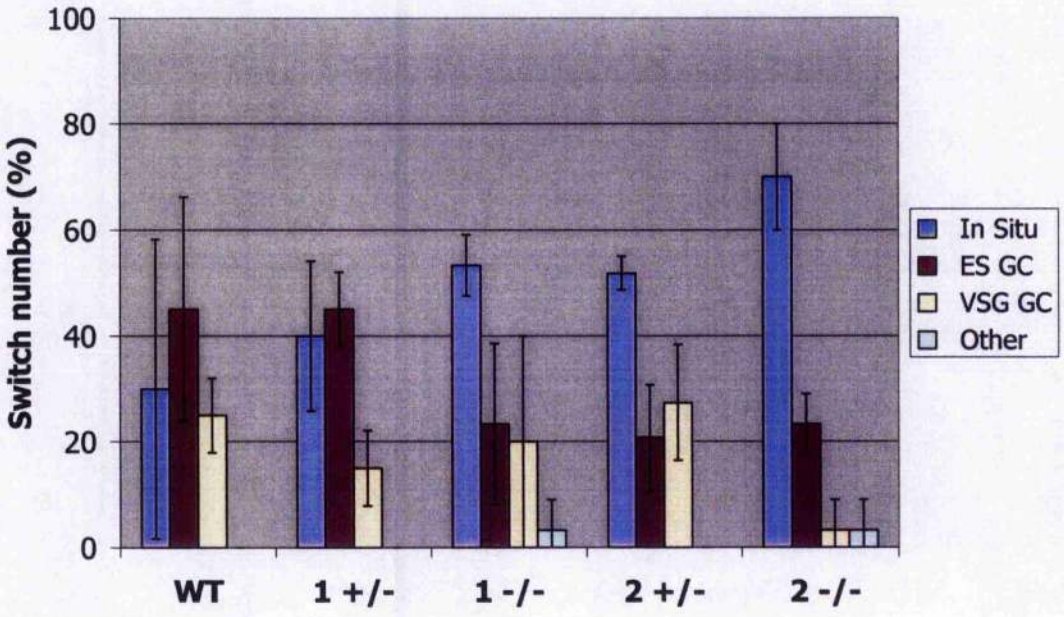


Figure 4.32: Analysis of switching events in the *DMC1* mutants. In the table (top) the type of switching mechanism used in a number of putatively clonal switched variants, from each of the *DMC1* cell lines and 3174.2 wild type cells, is indicated. The data are shown for the individual experiments detailed in the text. The graph (bottom) shows the tabulated data averaged for each cell line, and standard deviations are indicated. For both parts, the following nomenclature is used: In situ: in situ transcriptional switch. ES GC: expression site gene conversion. VSG GC: VSG gene conversion. Other: unknown. WT: wild type. +/-: heterozygous mutant. -/-: homozygous mutant.

4.9 Summary

The aim of this chapter was to determine if *RAD51-3*, *RAD51-5* or *DMC1* had any roles in DNA damage repair, recombination or VSG switching in 427 bloodstream form *T. brucei*. To do this, a reverse genetics approach was taken and homozygous mutants generated for each gene. All three genes could be made mutant, indicating that none of the genes are essential in this life-cycle stage, as had been found for *RAD51* (McCulloch and Barry, 1999). Mutation of two of the genes, *RAD51-3* and *RAD51-5*, caused a growth defect, observed by an increased population doubling time *in vitro* and *in vivo*. The cause of this is unknown, but is not explained by arrest of the cells at a defined stage of the cell cycle. The gene that showed no growth defect on mutation, *DMC1*, displayed no phenotypes consistent with the encoded protein contributing the DNA damage repair, recombination or *VSG* switching, at least in bloodstream form cells. *RAD51-3* and *RAD51-5* in contrast, were shown to be impaired in their ability to repair MMS induced DNA damage and to integrate transformed DNA constructs into their genome. Both phenotypes are consistent with the encoded proteins contributing to *T. brucei* DNA repair and homologous recombination. Perhaps the most surprising finding, given the similar roles for RAD51-3 and RAD51-5 in *T. brucei* general repair and recombination, is that *RAD51-3* mutants were found to be impaired in *VSG* switching, while *RAD51-5* mutants were not. The *RAD51-3* mutants, like *RAD51*, appeared to have an equal reduction in both *in situ* transcriptional switching and gene conversion switching.

CHAPTER 5

THE INTERACTIONS OF *RAD51*-LIKE PROTEINS IN *T. BRUCEI*

5.1 Introduction

The presence of DSBs has been shown in a number of organisms to cause cell cycle checkpoint arrest and the visible relocalisation of recombination enzymes to form putative repair complexes known as foci (Gasior *et al.*, 1998; Haaf *et al.*, 1995; Haaf *et al.*, 1999; Lisby *et al.*, 2001; Lisby *et al.*, 2003a; Lisby *et al.*, 2003b; Lisby *et al.*, 2004; Lisby and Rothstein, 2004; Lukas *et al.*, 2003; Maser *et al.*, 1997; Melo *et al.*, 2001; Raderschall *et al.*, 1999). The formation of foci was first observed in meiotic yeast cells in which Rad51 and Dmc1 were shown to co-localise in what was postulated to be recombination intermediates (Bishop, 1994). It was later shown that Rad52 and RPA foci also co-localise with Rad51 during meiosis, and that RPA foci formed during late G_1 and S phases (Gasior *et al.*, 1998). Yeast repair foci are also formed in mitotic cells, either as spontaneous events or in response to induced DNA damage. Here, the formation of Rad51 and Rad52 foci is limited to the S and G_2 phases, suggesting that the repair of the DSBs are regulated in a cell cycle dependent manner (Lisby *et al.*, 2003a; Lisby *et al.*, 2004). The observation that Rfa1 foci (presumably as the RPA complex) are formed in G_1 cells (Lisby *et al.*, 2004) indicates that DSBs arise and can be processed at any stage of the cell cycle, but the repair appears to be cell cycle limited. Rad52, Rad55 and Rad57 are essential for the formation of Rad51 foci in yeast meiotic cells. In contrast, only Rad52 appears to be involved in the formation of DNA damage-induced Rad51 foci (Gasior *et al.*, 1998; Lisby *et al.*, 2004), and this response to DNA damage is combined with an increase in the expression of Rad51 (Mercier *et al.*, 2001). Analysis of yeast Rad52 sub-nuclear localisation showed that the formation of spontaneous foci is almost exclusively limited to budded cells, suggesting that they are formed due to DNA damage generated during replication (Lisby *et al.*, 2001; Lisby *et al.*, 2003b). In contrast, exogenous generation of DSBs by exposure to γ-irradiation resulted in a dramatic increase in the numbers of cells containing foci (Lisby *et al.*, 2001). Together, these data indicate that repair foci are responsible for repairing DSBs generated in a number of ways. Rad52 foci were also shown to contain multiple DSBs, as the number of foci present in yeast cells is always lower than the number of DSBs generated (Lisby *et al.*, 2003b). This led the authors to postulate that the aggregation of recombination and checkpoint proteins at a few repair centres may help facilitate the coordination of the DNA damage response.

As in yeast, the formation of Rad51 foci in humans (Haaf *et al.*, 1995), and Rad51 or Rad52 foci formation in mice (Liu and Maizels, 2000), has been demonstrated to be

dramatically increased in response to DNA damage. In contrast to yeast, however, the formation of Rad51 foci in response to DNA damage in CHO cells was not combined with increased expression of Rad51 (Bishop *et al.*, 1998; Mercier *et al.*, 2001). These data suggest that DNA repair foci observed in mammalian cells are purely a result of the redistribution of pre-existing protein. Further analysis of this redistribution showed that Rad51 formed foci at sites of single strand DNA (Raderschall *et al.*, 1999).

In mammalian cells, as in yeast meiosis, loss of the Rad51 paralogues can have deleterious effects on the formation of Rad51 foci. CHO cell lines deficient in the Rad51 paralogues Xrcc2 and Xrcc3 were shown to have a deficiencies in Rad51 foci formation that were recovered by the transformation of plasmids expressing either Xrcc2 or Xrcc3 (Bishop *et al.*, 1998; O'Regan *et al.*, 2001). It was later shown that Xrcc3 is recruited to the sites of DSBs at an early stage of the repair process and independently of Rad51 (Forget *et al.*, 2004). Takata et al (2001) shows that Rad51C, Rad51D, Xrcc2 and Xrcc3 also contribute to Rad51 foci formation in chicken DT40 cells. It therefore appears that Xrcc3 in mammalian cells may be acting in a similar manner to Rad52 in yeast, in that it acts early in the DNA damage response. It appears that vertebrate Rad51-like proteins, unlike Rad55 and Rad57 in yeast, act in Rad51 foci formed mitotic cells. In mice the formation of Rad51 foci was observed to occur at a rate slower than that of Rad52 foci (Liu and Maizels, 2000). In fact, analysis of the diffusion rates of Rad51, Rad52 and Rad54 in mammalian nuclei showed all three to have different mobilities (Essers *et al.*, 2002), suggesting that they are not components of a preassembled DNA repair holo-enzyme in the absence of DNA damage, but are able to diffuse freely within the cell. This appears to be confirmed by the finding that Rad51, Rad52 and Rad54 are able to diffuse into or out of foci formed in response to DNA damage (Essers *et al.*, 2002).

In keeping with the apparent distinction between spontaneous and induced Rad52 foci and yeast, the formation of mammalian Rad51 foci differs depending on the source of DNA damage. Human cells containing a truncation of the breast cancer tumour suppressor gene *Brca2* display an absence of Rad51 foci in response to γ-irradiation. However, similar levels of S-phase, or spontaneous, Rad51 foci formation is observed in the same cells before and after exposure to γ-irradiation (Tarsounas *et al.*, 2003; Tarsounas *et al.*, 2004a). These results suggest that there are distinct pathways for the assembly of RAD51 foci in response to stalled replication forks or to irradiation.

So far this project has analysed the roles of *T. brucei* RAD51-3, RAD51-5 and DMC1 on an individual basis. The aim of this was to determine which of the three proteins had

roles in DNA repair, recombination and antigenic variation. Now that roles have been identified for RAD51-3 and RAD51-5 in repair and recombination, and for RAD51-3 in antigenic variation, the question arises: do these proteins interact with RAD51? The aim of this chapter is to start to determine the interactions, if any, between the *RAD51*-like genes. This question was addressed in two ways. Firstly, *in situ* hybridisation assays were carried out to assess the ability of each of the *RAD51*-like gene mutants to influence the formation of RAD51 foci. Secondly, attempts were made to over-express each of the *RAD51*-like genes in a *rad51-/-* mutant background. This was intended to examine if any of the RAD51-like proteins can compensate for the loss of RAD51, especially in response to DNA damage, indicating that they may be capable of catalysing DNA strand exchange independently of RAD51.

5.2 *in situ* hybridisation

As a means to examine the function of *T. brucei* RAD51-3 and RAD51-5 in the regulation of RAD51 expression, this chapter exploits recent unpublished work by Katherine Norrby, which demonstrates that *T. brucei* RAD51 localises in sub-nuclear foci following damage. Since the work I will describe is reliant on these findings of K. Norrby, a summary of this analysis is provided in the following section to clarify my later work.

To analyse whether or not *T. brucei* forms RAD51 foci in response to DNA damage as has been described in yeast and mammals, *in situ* hybridisation was carried out on Lister 427 wild type bloodstream form cells using rabbit polyclonal antisera raised against *T. brucei* RAD51 expressed in *E. coli* (referred to as 'anti-RAD51 antiserum') and secondary hybridisation using a goat derived anti-rabbit IgG with FITC conjugate (Sigma; referred to as 'FITC conjugate').

The cells were grown, from a starting density of 1 x 10^6 cells.ml^{-1}, in HMI-9 containing a range of phleomycin concentrations for 18 hours and then examined under a microscope for the presence or absence of foci (section 2.14.2; details of this analysis are described in Section 5.2.2). The results show that wild type cells in the absence of DNA damage showed predominantly RAD51 staining throughout the cell (Fig 5.1), with no evidence for nuclear localisation. It seems likely that this cytoplasmic localisation of RAD51 is genuine and not due to substantial cross reaction of the antiserum, since very little staining was seen in *rad51-/-* cells.

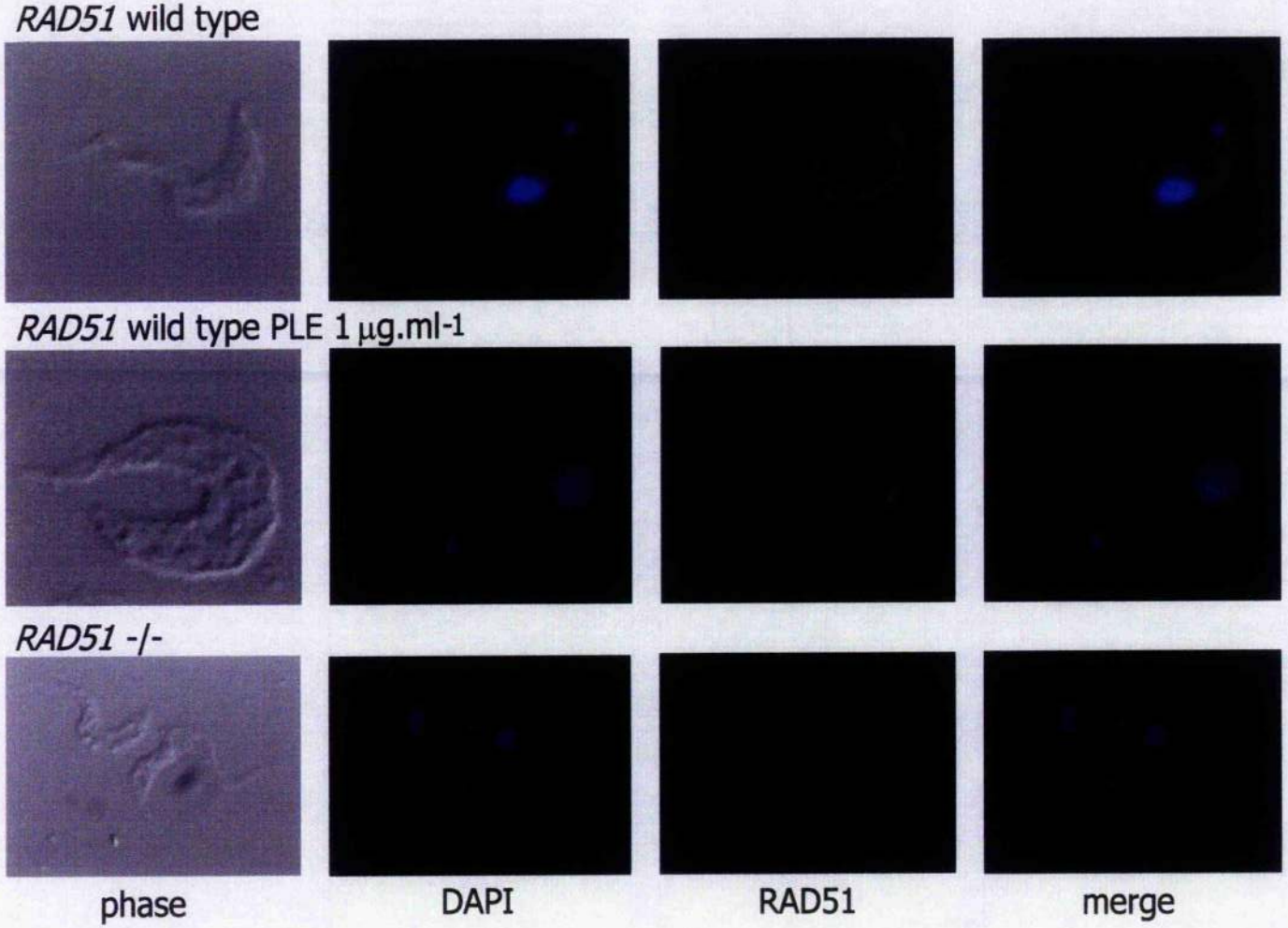


Figure 5.1: RAD51 sub-cellular localisation in wild type and *RAD51*-/- cells. Each cell or group of cells is shown in phase contrast (phase), after staining with DAPI (DAPI) or after hybridisation with anti-RAD51 antiserum and secondary hybridisation with FITC conjugate (RAD51). Merged images of DAPI and FITC stained cells are also shown.

Western analysis carried out using total protein extracts from wild type and *rad51-/-* cells showed that the anti-RAD51 antiserum recognises primarily RAD51 in bloodstream form cells, since there was very limited hybridisation in *rad51-/-* extract (Fig 5.2): Some minor cross-reaction with an unidentified protein species of around 80 kDa was apparent. In response to DNA damage, RAD51 was observed to co-localise with the nucleus and to form foci. Increasing the phleomycin concentration, up to 5 $\mu g.ml^{-1}$, resulted in increased numbers of cells with sub-nuclear foci and an increase in the number of foci within those cells (Table 5.1). At 0.2 and 0.5 $\mu g.ml^{-1}$ phleomycin 42.2 – 48.1% of cells formed 1 or more RAD51 foci, with approximately 27 – 32% of cells containing 2 – 4 foci. Increasing the concentration of phleomycin to 1.0 and 2.0 $\mu g.ml^{-1}$ resulted in 74 – 87.7% of cells forming 1 or more RAD51 foci, with approximately 45 – 50% of cells containing 2 – 4 foci. However, increasing the concentration further, to 5.0 $\mu g.ml^{-1}$ of phleomycin, resulted in undetectable RAD51 expression, presumably because the level of DNA damage is so great that it has started to affect transcription.

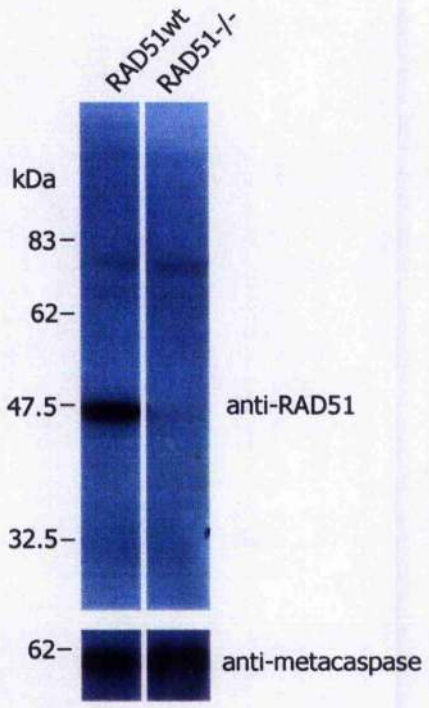


Figure 5.2: Western of wild type and *rad51-/-* bloodstream *T. brucei* probed with the anti-RAD51 antiserum. Total protein extracts from wild type Lister 427 bloodstream form cells and from 427 rad51-/- bloodstream form cells is shown after hybridisation with anti-RAD51 antiserum. Size markers are indicated and a loading control with antisera against metacaspase5 (J. Mottram and M. Helms, gift) is shown.

		Number of foci (%)							
	[PLE]	0	1	2	3	4	5	6	M
RAD51 wild type	0	99.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
	0.2	57.8	14.7	14.7	10.1	1.8	0.9	0.0	0.0
	0.5	51.9	13.5	13.5	5.8	12.5	1.0	1.0	1.0
	1	26.0	16.3	16.3	11.5	17.3	5.8	1.0	5.8
	2	12.3	14.2	14.2	32.1	3.8	0.9	0.0	22.6
RAD51 -/-	0	99.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0

Table 5.1: RAD51 foci formation in wild type and *rad51*-/- 427 bloodstream form cells. The percentage of cells containing RAD51 foci, the numbers of foci are shown at varying concentrations of phleomycin (PLE). Concentrations of phleomycin are in $\mu g.ml^{-1}$. M: more than 6 foci. -/-: homozygous mutant.

It is possible that the RAD51 foci that were observed may simply result from an increase in expression of RAD51 in response to DNA damage, resulting in punctuate staining in the nucleus, rather than re-localisation of RAD51 to the nucleus. I have already shown that *RAD51* mRNA does not increase in response to DNA damage (section 3.8), but to examine if the same is true for the RAD51 protein, a Western was carried out (by Richard McCulloch) on total protein from 427 bloodstream form cells exposed to increasing concentrations of phleomycin (Fig. 5.3). These results indicate that RAD51 protein levels also do not increase in response to DNA damage.

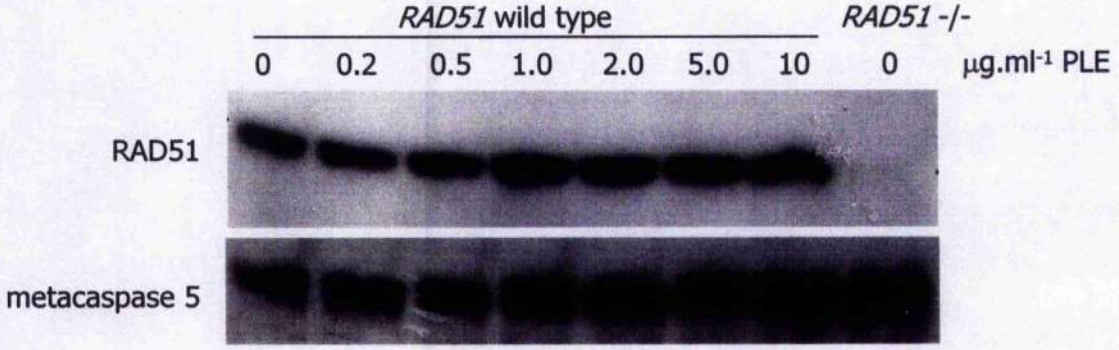


Figure 5.3: Western of protein extracts from cells exposed to phleomycin induced DNA damage. Lister 427 bloodstream form cells were exposed to increasing concentrations of phleomycin (PLE) as indicated and total protein extracts probed with anti-RAD51 antiserum; anti-metacaspase 5 antiserum was used as a loading control.

This work demonstrates that in situ hybridisation analysis is a viable approach to assess whether or not *RAD51-3* and *RAD51-5* have roles in RAD51 function. By exposing *rad51-3* and *rad51-5* homozygous mutant cell lines to DNA damage, the contribution of the protein products to the formation of foci can be determined. This would address directly the question of whether or not RAD51-3 and RAD51-5 operate in the same

DNA damage repair pathway as RAD51. Given that the DMC1 mutants have no discernable defects in repair or recombination function, at least in bloodstream form cells, the contribution of *DMC1* to the formation of RAD51 foci was not examined.

5.2.1 Assessing the phleomycin sensitivity of *RAD51-3* and *RAD51-5* mutants

It has already been shown that *rad51-3* and *rad51-5* homozygous mutants have increased sensitivity to the alkylating agent MMS (section 4.5). To examine if this DNA damage phenotype is also the case for phleomycin a simpler sensitivity assay was carried out. Wild type cells and *RAD51-3* or *RAD51-5* heterozygous and homozygous mutant cell lines were plated out at 10^6 cells.ml^{-1}, in a total volume of 1 ml, in varying concentrations of phleomycin and cell counts carried out at 18 and 42 hour time points (Table 5.2). It appears that over the time course of this assay, the expected increase in phleomycin sensitivity of the *rad51-3* and *rad51-5* homozygous mutants relative to the wild type cells and heterozygous mutants was not discernable. In all cases, increasing the concentration of phleomycin showed an increasing extent of cell death and the amount of increased cell death or growth retardation after 42 hrs compared with 18 hours are comparable. The lower cell concentrations of the homozygous mutants observed in all conditions are most readily explained by their longer generation times (Section 4.3.1). It is possible, therefore, that the *rad51-3* and *rad51-5* homozygous mutants are no more sensitive to the DNA damage caused by phleomycin and show a distinction in this form of DNA repair relative to that which is required after exposure to MMS. Indeed, such a situation has also been seen previously in *T. brucei*, since *mre11-/-* mutants do not display increased sensitivity to MMS, relative to wild type cells and heterozygous mutants, but are more sensitive to phleomycin (Robinson *et al.*, 2002). Nevertheless, the assay used in this work is not the same as the MMS survival assay used previously (Section 4.5) or in the *MRE11* analysis (Robinson *et al.*, 2002), as it uses higher concentrations of damaging agent and scores the amount of cell death rather than the ability of individual cells to survive the treatment. For this reason, we have assumed that the *rad51-3* and *rad51-5* homozygous mutants do display increased phleomycin sensitivity in examining RAD51 foci formation and examined foci formation at 2 phleomycin concentrations.

		Cell line									
		WT		**1 +/-**		**1 -/-**		**2 +/-**		**2 -/-**	
	Time (hours)	18	42	18	42	18	42	18	42	18	42
	[Ple] µg.ml^{-1}										
RAD51-3	0.1	X	X	X	X	0.90	1.12	X	X	1.00	1.60
	0.3	X	X	X	X	0.92	0.58	X	X	0.88	0.37
	0.6	1.64	1.04	1.60	0.70	0.64	0.25	2.08	1.08	0.54	0.14
	0.9	1.16	0.76	1.20	0.44	0.67	0.23	0.92	0.78	0.53	0.18
	1.2	0.88	0.36	0.70	0.34	0.82	0.31	1.48	0.39	0.68	0.19
	1.5	1.04	0.29	0.82	0.33	0.88	0.17	1.12	0.40	0.46	0.14
	1.8	1.16	0.36	0.90	0.23	0.64	0.25	0.72	0.15	0.62	0.10
	2.1	0.88	0.34	0.92	0.32	0.52	0.12	1.04	0.35	0.36	0.16
	2.4	1.24	0.32	0.84	0.21	0.53	0.23	1.00	0.16	0.48	0.12
RAD51-5	0.1	X	X	X	X	1.40	2.20	X	X	1.36	0.45
	0.3	X	X	X	X	1.18	0.65	X	X	1.40	0.36
	0.6	0.92	0.51	1.27	0.90	0.93	0.38	1.35	1.04	1.04	0.25
	0.9	0.88	0.38	0.95	0.63	0.96	0.37	1.04	0.56	0.90	0.40
	1.2	0.80	0.32	0.84	0.51	0.90	0.33	0.98	0.39	0.93	0.31
	1.5	0.82	0.29	0.83	0.38	0.84	0.30	0.90	0.42	0.82	0.18
	1.8	0.84	0.36	0.65	0.36	0.75	0.21	0.78	0.40	0.76	0.12
	2.1	0.70	0.41	0.75	0.34	0.76	0.19	0.85	0.43	0.80	0.11
	2.4	0.62	0.18	0.80	0.42	0.78	0.21	0.82	0.44	0.78	0.24

Table 5.2: Phleomycin sensitivity assay for *RAD51-3* and *RAD51-5* mutant cell lines. Cell counts (at 10^6.ml^{-1}) are shown at varying concentrations of phleomycin after 18 and 42 hours growth. [Ple]: phleomycin concentration. X: experiment not carried out. WT: wild type. +/-: heterozygous mutant. -/-: homozygous mutant.

Firstly, *in situ* hybridisation analysis, for all of the cell lines, was conducted at a phleomycin concentration of 1.5 $\mu g.ml^{-1}$. This concentration resulted in growth arrest, or slight population death, in all of the cell lines and is in the range known to induce RAD51 foci formation in wild type cells (section 5.2). Secondly, to ensure that the *rad51-3* and *rad51-5* homozygous mutants were not damaged beyond the point at which RAD51 foci form, lower drug concentrations were also used: 0.3 $\mu g.ml^{-1}$ for *rad51-3-/-* and 0.5 $\mu g.ml^{-1}$ for *rad51-5-/-*. These concentrations gave comparable levels of cell death or growth retardation relative to the wild type cell and heterozygous mutants. Presumably a higher concentration is needed for *rad51-5-/-* mutants, as they do not have such a severe growth defect as the *rad51-3-/-* mutants (Section 4.3.1).

5.2.2 *in situ* hybridisation analysis of RAD51 in *RAD51-3* and *RAD51-5* mutants.

To examine the localisation of RAD51 following phleomycin damage, 5 ml cultures of all cell lines were set up at 10^6 $cells.ml^{-1}$ at 1.5 $\mu g.ml^{-1}$ phleomycin and additional cultures were set up for *rad51-3-/-* cells at 0.3 $\mu g.ml^{-1}$ and *rad51-5-/-* cells at 0.5 $\mu g.ml^{-1}$. After 18 hours growth, 1.5 mls of cells were spun down at 1500g for 10 mins, resuspended in 1ml of PBS to wash out residual HMI-9, spun down again and resuspended in 40 µl of PBS. 20 µl samples were then smeared on a microscope slide and allowed to air dry before being fixed and permeabilised by soaking in methanol for 10 mins. The slides were then washed in PBS containing 0.2% Tween-20 (PBS-T) for 5 mins before being moved to a humid chamber. 1ml of anti-RAD51 antiserum, diluted 1:500 in PBS containing 1% Tween-20 and 3% BSA (PBS-T-BSA), was added to the slide and left to hybridise for 90 mins at room temperature. The slides were rinsed with PBS-T-BSA and soaked twice in fresh batches of PBS-T-BSA for 5 mins before being returned to the humid chamber, at which point 0.5 ml of the FITC conjugated goat-derived anti-rabbit IgG, diluted 1:100 in PBS-T-BSA, was added to the slide and left in the dark to hybridise for 30 mins at room temperature. The slides were then rinsed with PBS-T-BSA and soaked twice in fresh batches of PBS-T for 5 mins before one drop of vectashield with DAPI (Vector laboratorics) was placed on the slide, a coverslip added and sealed with nail varnish. Slides were visualised the same day and counts of 100-120 cells were carried out and scored for the number of sub-nuclear foci that were visible. From these data, the percentages of each cell line containing 0 – 6, or more (M), discernable foci were calculated (Tables 5.3 and 5.4).

		Number of foci (%)							
	[PLE]	0	1	2	3	4	5	6	M
Wild Type	1.5	35.0	28.2	21.4	10.3	3.4	3.4	0.9	0.9
RAD51-3 +/- 1	1.5	27.4	30.8	18.8	15.4	6.0	1.7	0.0	0.0
RAD51-3 -/- 1	0.3	81.7	7.3	5.5	5.5	0.0	0.0	0.0	0.0
RAD51-3 -/- 1	1.5	63.8	18.1	12.1	3.4	2.6	0.0	0.0	0.0
RAD51-3 +/- 2	1.5	22.2	26.5	18.8	14.5	4.3	1.7	0.9	1.7
RAD51-3 -/- 2	0.3	59.0	19.7	12.0	4.3	0.9	0.9	0.0	0.0
RAD51-3 -/- 2	1.5	52.1	24.8	8.5	0.9	1.7	0.9	0.0	0.9

Table 5.3: RAD51 foci formation in wild type cells and *RAD51-3* mutants. Percentages of cells that show foci and the number of foci at given concentrations of phleomycin (PLE) are shown. Concentrations of phleomycin are in $\mu g.ml^{-1}$. M: more than 6 foci. +/-: heterozygous mutant. -/-: homozygous mutant.

		Number of foci (%)							
	[PLE]	0	1	2	3	4	5	6	M
Wild Type	1.5	22.3	16.1	24.1	17.9	11.6	4.5	1.8	1.8
RAD51-5 +/- 1	1.5	19.8	24.3	33.3	14.4	5.4	1.8	0.9	0.0
RAD51-5 -/- 1	0.5	67.3	18.3	16.3	5.8	1.0	0.0	0.0	0.0
RAD51-5 -/- 1	1.5	59.5	18.0	9.0	2.7	0.9	0.9	0.0	0.9
RAD51-5 +/- 2	1.5	27.9	32.4	19.8	13.5	2.7	4.5	3.6	3.6
RAD51-5 -/- 2	0.5	64.0	22.5	7.2	3.6	0.9	0.0	0.9	0.9
RAD51-5 -/- 2	1.5	64.9	23.4	3.6	1.8	2.7	0.0	0.0	0.0

Table 5.4: RAD51 foci formation in wild type cells and *RAD51-5* mutants. Percentages of cells that show foci and the number of foci at given concentrations of phleomycin (PLE) are shown. Concentrations of phleomycin are in $\mu g.ml^{-1}$. M: more than 6 foci. +/-: heterozygous mutant. -/-: homozygous mutant.

From these results it can be seen that loss of functional RAD51-3 and RAD51-5 results in a reduced capacity for RAD51 foci formation. A reduction in the number of cells with visible RAD51 foci, and a concomitant increase in the number of cells without foci were apparent in the *rad51-3-/-* and *rad51-5-/-* mutants relative to the wild type and heterozygous cells. This was true at either the lower or higher concentrations of phleomycin used in the homozygous mutants. Moreover, the number of foci in the homozygous mutant cells was also reduced: only very small numbers of cells were seen with greater than 3 - 4 foci (Tables 5.3 and 5.4).

The RAD51 foci that formed in wild type cells were generally clear and distinct, with the majority of cells containing 1 – 3 foci (examples are shown Fig. 5.4). In contrast, while, the *RAD51-3* and *RAD51-5* heterozygous cell lines exhibited qualitatively the same RAD51 foci as wild type cells, the foci in the homozygous cell lines were clearly different (Fig. 5.5). The foci formed in the *rad51-5-/-* cell lines, in general, were much less well-defined with a significant level of cytoplasmic RAD51 staining remaining (Figs. 5.5 and 5.6). Indeed, in many cases (such as in Fig. 5.6D) the level of cytoplasmic RAD51 staining was such that the foci observed were exceptionally faint. In the *rad51-*

3-/- cell lines, these observations were also apparent, although the cells tended to show less cytoplasmic localisation and more distinct foci (e.g. Fig. 5.7), though still not as clearly as in the wild type or *RAD51-3+/-* cells.

These results suggest that both proteins facilitate the movement of RAD51 into sub-nuclear foci following phleomycin-induced DNA damage. The increased cytoplasmic staining in the *rad51-3-/-* and *rad51-5-/-* mutants may suggest, furthermore, that RAD51 is transferred from the cytoplasmic in to the nucleus in this reaction, either by RAD51-3 or RAD51-5 directly, or by other factors. We cannot exclude however, that the cytoplasmic staining results from cross reaction of the anti-RAD51 antiserum with another unidentified protein and that this is more apparent when trying to visualise the fainter sub-nuclear RAD51 foci in the homozygous mutants.

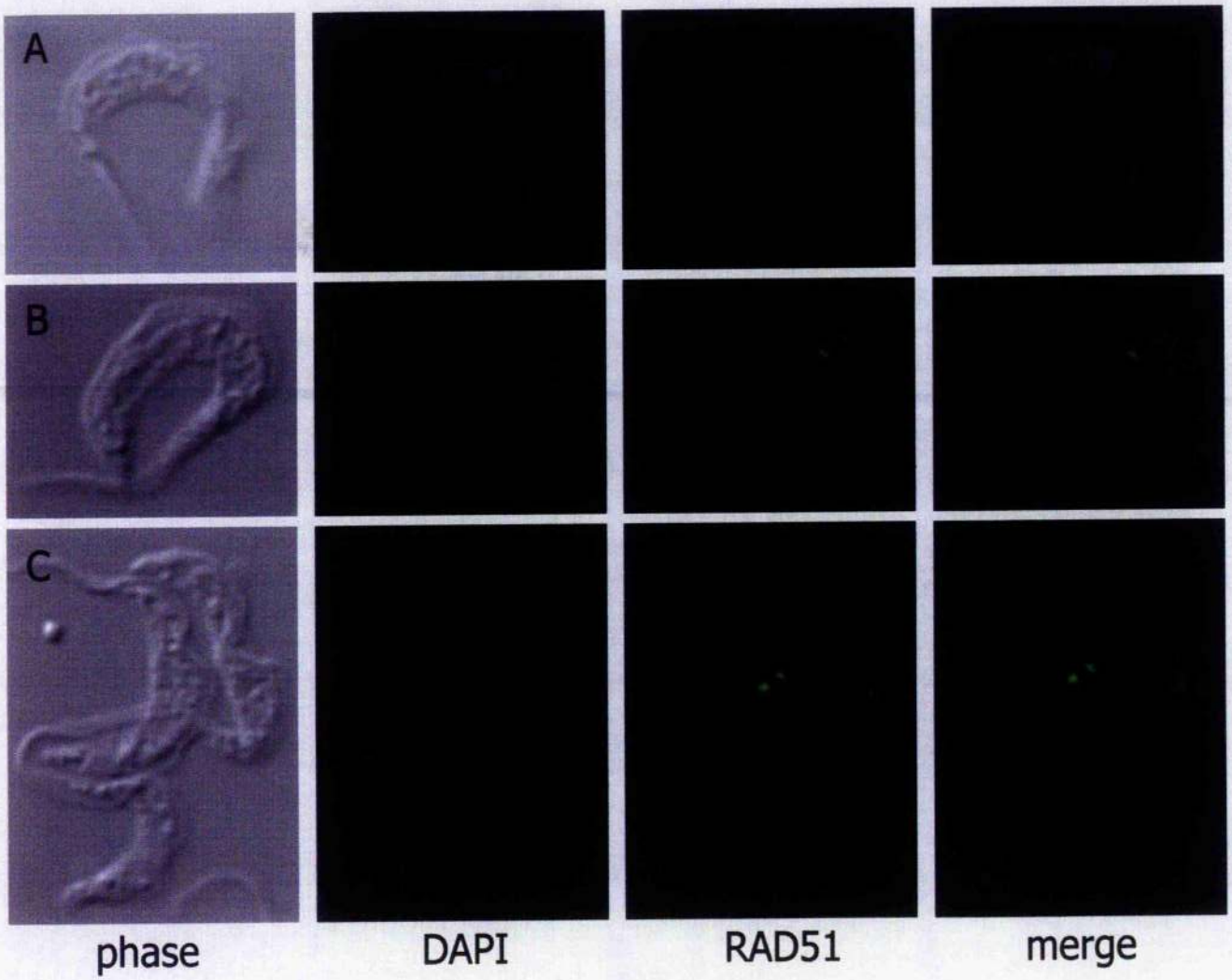


Figure 5.4: RAD51 foci formation in wild type cells. Each cell or group of cells is shown in phase contrast (phase), after staining with DAPI (DAPI) or after hybridisation with anti-RAD51 antiserum and secondary hybridisation with FITC conjugate (RAD51). Merged images of DAPI and FITC stained cells are also shown. A: a cell with 8 foci, B: a cell with 3 foci. C: cells with 0, 1 and 2 foci.

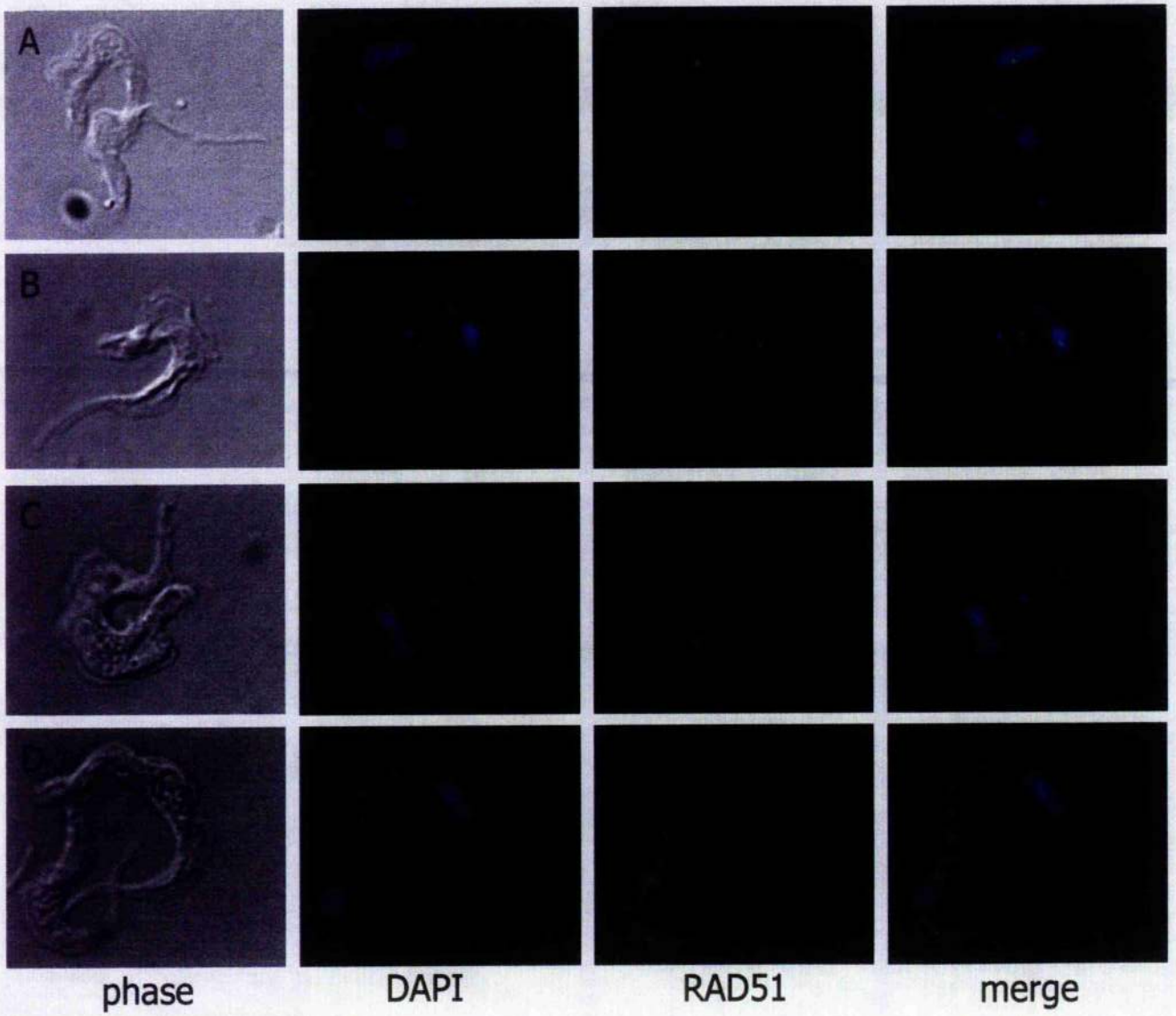


Figure 5.5: RAD51 foci formation in *RAD51-3* and *RAD51-5* heterozygous and homozygous cell lines. Each cell or group of cells is shown in phase contrast (phase), after staining with DAPI (DAPI) or after hybridisation with anti-RAD51 antiserum and secondary hybridisation with FITC conjugate (RAD51). Merged images of DAPI and FITC stained cells are shown also. All cells were exposed to 1.5 $\mu g.ml^{-1}$ phleomycin. A: *RAD51-3* heterozygous cells, one with a single focus. B: a *RAD51-3* homozygote cell with a single focus. C: a *RAD51-5* heterozygote cell with a single focus. D: *RAD51-5* homozygous cells, one with a single focus.

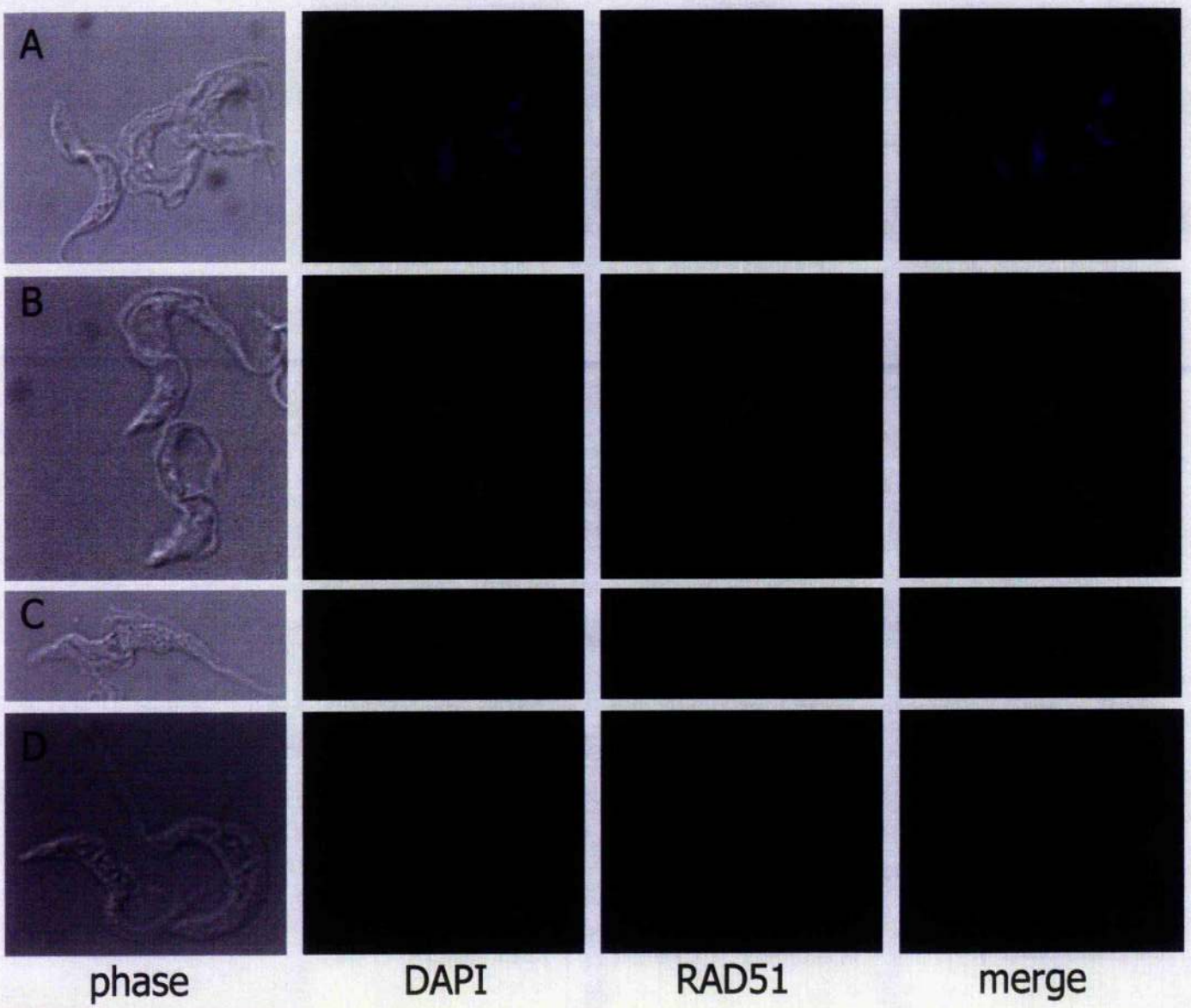


Figure 5.6: RAD51 foci formation in *RAD51-5* homozygous cell lines. Each cell or group of cells is shown in phase contrast (phase), after staining with DAPI (DAPI) or after hybridisation with anti-RAD51 antiserum and secondary hybridisation with FITC conjugate (RAD51). Merged images of DAPI and FITC stained cells are shown also. A: three cells each with a single focus, B: two cells, the top one contains 2 foci, C: a cell with a single focus and D: two cells both with 2 foci.

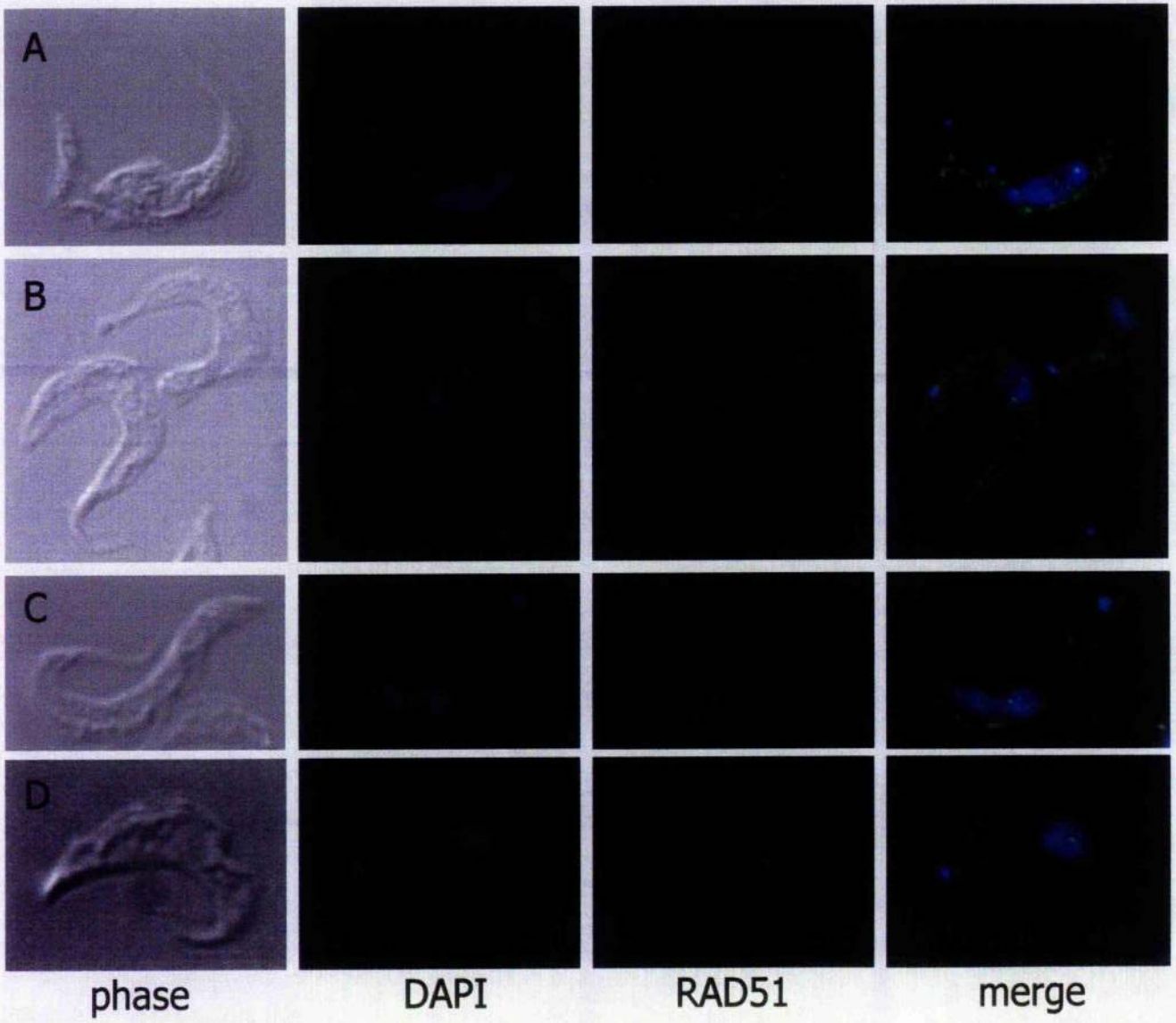


Figure 5.7: RAD51 foci formation in *RAD51-3* homozygous cell lines. Each cell or group of cells is shown in phase contrast (phase), after staining with DAPI (DAPI) or after hybridisation with anti-RAD51 antiserum and secondary hybridisation with FITC conjugate (RAD51). Merged images of DAPI and FITC stained cells are shown also. All cells were exposed to 1.5 $\mu g.ml^{-1}$ phleomycin. A: a cell with a single focus, B: two cells both with 2 foci, C: a cell with 2 foci and D: a cell with a single focus.

5.3 Complementation of *rad51 -/-*

rad51-/- cells have been shown to have DNA damage sensitivity, a recombination defect and a reduced level of VSG switching. However, all of these processes do still occur, albeit at a reduced level. This suggests that other enzymes are capable of catalysing homologous recombination in the absence of RAD51. In an attempt to determine if RAD51-3, RAD51-5 or DMC1 can fulfil this function, attempts were made to over-express each gene in *rad51-/-* cells. The over-expression cell lines could then be analysed for any reduction in the defects caused by the loss of RAD51, indicating that the RAD51-like proteins can operate as independent enzymes in homologous recombination.

To attempt this complementation of *rad51-/-* mutants, the inducible expression system comprising the plasmids pLew13 (Wirtz *et al.*, 1999) and pHD789 (Irmer and Clayton, 2001) kindly donated by C. Clayton (Fig. 5.8) was chosen. This system, which utilises T7 RNA polymerase and the tetracycline repressor/operator system, allows inducible over-expression from any cloned gene. The construct pLew13 expresses the tetracycline repressor, T7 RNA polymerase and neomycin phosphotransferase (providing selection of transformants that have integrated the plasmid into the *T. brucei* tubulin array following electroporation). Integration into the tubulin array results in constitutive expression of the tetracycline repressor and T7 RNA polymerase. pLew13 transformants are next transformed with pHD789, a construct which integrates in to the rRNA spacer and contains the T7 RNA polymerase promoter, the tetracycline operator and the gene to be over-expressed. pHD789 transformants are selected for using a hygromycin resistance marker. When double transformant cells are grown in the absence of tetracycline, the repressor binds to the operator, blocking transcription from the T7 promoter. Addition of tetracycline to the culture media results in a release from the transcriptional block, thereby allowing T7 polymerase to transcribe any gene of interest.

5.3.1 Attempted generation of *RAD51-3*, *RAD51-5* and *DMC1* over-expression cell lines

The pLew13 construct was linearised by restriction digestion with *Not*I and the DNA then phenol:chloroform extracted and ethanol precipitated. 5 µg of digested plasmid

was used in each of three transformations of 5 x 10^7 *rad51* -/- cells, carried out as described in Materials and Methods Section 2.4. Antibiotic resistant transformants were first selected using 2.5 $\mu g.ml^{-1}$ of G418 and 1 x 10^7 cells were plated out over 12 wells using 1.5 mls HMI-9 per well. Since no transformants were obtained from this approach, the experiment was repeated in exactly the same way except that a slightly reduced level of G418 (2 $\mu g.ml^{-1}$) was used for selection. Following this approach, a total of 3 transformants were generated. The low frequency of transformation is almost certainly a consequence of the *rad51* mutation, since a transformation efficiency approximately 4-fold greater was seen with pLew13 in wild type cells (data not shown). Genomic DNA was extracted from all three putative *rad51*-/- pLew13 transformant cell lines and subjected to restriction digestion using *Not*I. The digested DNA was subsequently separated by gel electrophoresis in a 0.8% agarose 1 x TBE gel, Southern blotted and probed with a fragment of the neomycin resistance gene (generated by PCR using Taq DNA polymerase and primers Neo-5′ and Neo-3′; Appendix 1).

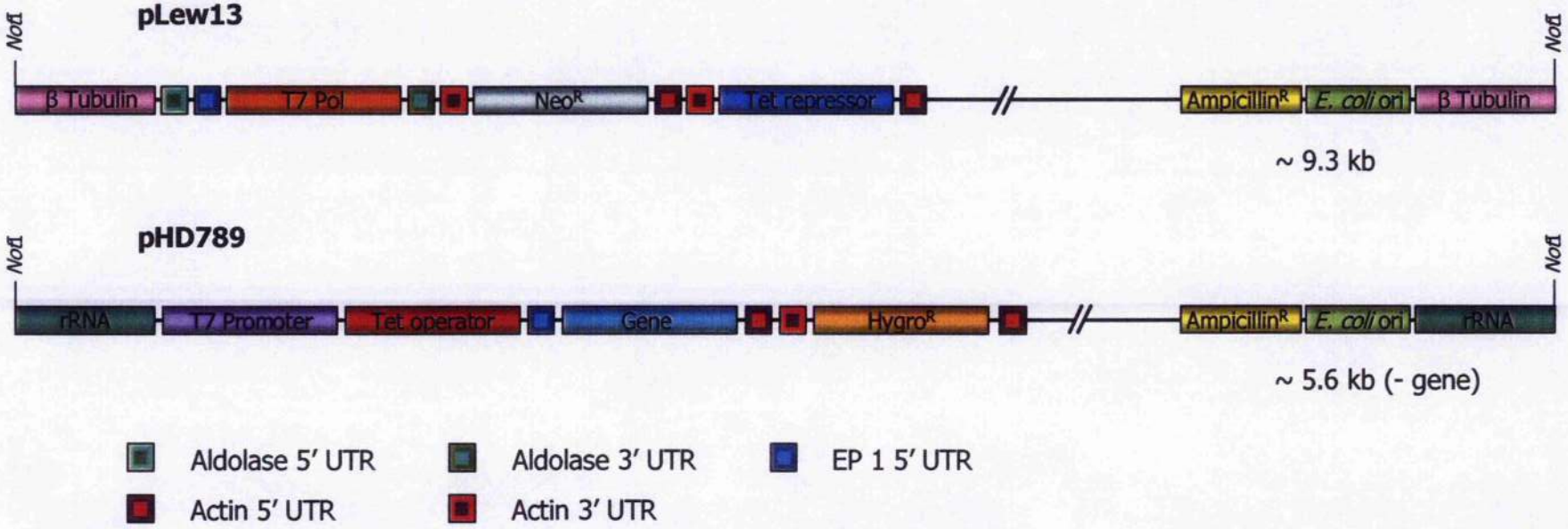


Figure 5.8: Gene over-expression constructs used in this study. Cells are first transformed with pLew13 (shown top), which integrates into the tubulin array and results in constitutive expression of T7 RNA Polymerase (T7 Pol) and the tetracycline (Tet) repressor. Transformants are selected for using the neomycin (Neo^R) resistance gene. PLew13 transformants are next transformed using the gene over expression construct pHD789, which integrates into the rRNA spacer (rRNA). This construct contains the T7 RNA polymerase promoter (T7 promoter), tetracycline (Tet) operator, the gene of interest (Gene) and transformants are selected for using the hygromycin ($Hygro^R$) resistance gene. Both plasmids are integrated as whole molecules, so transformants also contain the replication origin (*E. coli* ori) and the ampicillin resistance gene ($Ampicillin^R$) for growth and selection in *E. coli* respectively.

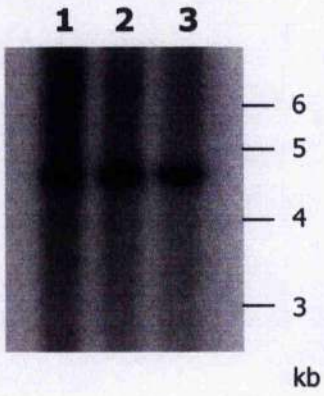


Figure 5.9: A Southern of three putative *rad51-/-* pLew13 transformants. Genomic DNA was digested with NotI and probed with a fragment of the neomycin resistance gene. Size markers (in kbp) are shown

Correct integration of the pLew13 construct should be seen *via* the neomycin gene-containing fragment of approximately 9.3 kb in size, since *Not*I should be recreated in the integrated β tubulin open reading frame (Fig. 5.8) and the plasmid precisely excised by restriction digestion. The fragment sizes, seen in figure 5.9, do not match those predicted, meaning that the integration of the pLew13 construct is anomolous, almost certainly a result of the inactivation of *RAD51*. Most likely, the transformants integrated part of the construct using internal sequence homology due to the numerous 5′ and 3′ processing signals present within the construct. The actin sequence annotated by the *T. brucei* genome sequencing initiative was checked to see if the neomycin resistance gene had replaced one of the actin genes using the 5′ and 3′ un-translated regions for homologous recombination. This had not happened, however, as replacement of one of the actin genes would result in a *Not*I digestion fragment of more than 20 kb, much larger than the 4.6-4.7 kb band observed.

Due to time limitations, further attempts at the correct integration of the pLew13 construct in rad51-/- cells and the subsequent transformations with pHD789 containing *RAD51-3*, *RAD51-5* and *DMC1* (which have been generated; data not shown) were not attempted. It is conceivable that this approach remains viable, however, despite the technical problems associated with generating the appropriate lines of *rad51-/-* cells.

5.4: Summary

The aim of this chapter was to try to determine whether or not RAD51-3 and/or RAD51-5 interacts with RAD51, now that roles in DNA damage repair and homologous recombination have been identified for these proteins. The first part of this chapter involved an examination of the formation of RAD51 foci in *T. brucei*. 427 bloodstream

form cells were examined for RAD51 sub-cellular localisation, before and after DNA damage, in *RAD51-3* and *RAD51-5* mutants. From this, it appears that both genes appear to be involved in the regulation of the movement of RAD51 to sub-nuclear foci. In both *RAD51-3* and *RAD51-5* mutants there was a reduction in the number of cells exhibiting foci and a reduction of the number of foci within those cells. In addition, in the cells that contained foci, a higher level of cytoplasmic RAD51 localisation was apparent, and the foci formed were less distinct than those formed in wild type cells. Whether or not RAD51-3 and RAD51-5 are involved in the same pathway as RAD51 is a question that cannot be answered by this analysis.

In the second part, attempts were made to over-express *RAD51-3*, *RAD51-5* and *DMC1* in order to determine if any of them could complement *rad51-/-*. Over-expression constructs were generated for all three genes and test transformations were successful using wild type cells. However, we were not able to generate pLew13 transformants in a *rad51* -/- background. As a result of this, the experiment was abandoned due to time constraints.

CHAPTER 6

DISCUSSION

6.1 Introduction

The overall aim of this project was to identify *RAD51*-like genes present in the *T. brucei* genome, and to examine their functions. Blast searches of the *T. brucei* genome database, using *S. cerevisiae* Rad51, *T. brucei* RAD51 and *E. coli* RecA sequences as queries, identified five putative *RAD51*-like genes in addition to the canonical, previously characterised RAD51 (McCulloch and Barry, 1999). This number of potential RAD51-related functions, which is approximately 50% greater than that found in *S. cerevisiae* (Krogh and Symington, 2004), is not a peculiarity of *T. brucei*, as orthologues of these genes can also be identified in *T. cruzi*. Interestingly, one of the putative *RAD51*-related genes, which we have named *RAD51-5* (below), could not be found in *L. major*.

The assignment of the five genes as being likely to encode functions in DNA recombination and repair is based upon the identification of sequence motifs that are characteristic of Rad51, RecA and RadA proteins in eukaryotes, bacteria and archae respectively (Brendel *et al.*, 1997; Pellegrini *et al.*, 2002; Shao and Grishin, 2000; Story *et al.*, 1993). For three of the genes, genetic analysis was conducted to assign roles in the repair of induced DNA damage, homologous recombination and antigenic variation in Lister 427 bloodstream stage *T. brucei*. From these analyses, the broad conclusions (which are discussed further below) are as follows: no discernable function could be ascribed to one gene that is highly related to *RAD51* (*DMC1*); two more distantly related genes (*RAD51-3* and *RAD51-5*) were shown to have clear roles in DNA repair and in promoting a RAD51 response to induced DNA damage; and surprisingly, *RAD51-3*, but not *RAD51-5*, had a detectable role in antigenic variation.

6.2 The role of *DMC1*

DMC1 was shown to have no detectable functions in 427 bloodstream form cells in the processes of DNA damage repair, homologous recombination and antigenic variation, despite that fact that the gene is expressed in this life cycle stage, to generate apparently mature mRNA (Section 3.8.1), and can be disrupted (Section 4.2). One possible explanation for this is that *DMC1* does have roles in these processes but acts as a minor co-factor or in a minor pathway, meaning that the assays used could not detect this. Another possibility is that *DMC1* functions in general recombination, and repair, but in

life-cycle stages that we have not analysed. It should be possible to analyse the potential functions of *DMC1* in procyclic cells, as these can be readily cultured, but RAD51 appears to play a central function in recombination in this stage (McCulloch and Barry, 1999). A final possibility is that *DMC1* plays no roles in general repair and recombination in monomorphic cells, but may assume a central role in pleomorphic cells. Pleomorphic cells display distinct features of cellular differentiation compared to monomorphic cells (generating non-dividing short stumpy forms in a density dependent manner; (Matthews *et al.*, 2004; Tyler *et al.*, 2001; Vassella *et al.*, 1997) and, notably, display elevated frequencies of antigenic variation (Turner and Barry, 1989). Perhaps, then, DMC1 assumes a more central function than RAD51 in catalysing VSG switching in pleomorphic cells. Indeed, preliminary work suggests that the impairment of VSG switching caused by *RAD51* disruption in monomorphic *T. brucei* cells is not found in the same mutants in the ILTat 1.2 cell line of pleomorphic cells (P. Burton, PhD thesis). To examine this, *DMC1* gene disruption could be carried out in this *T. brucei* cell line.

The most likely explanation for the lack of a discernable role for *DMC1* in Lister 427 bloodstream stage cells, however, is that it has a meiosis-specific function. Sequence comparisons and phylogenetic analysis suggests that *T. brucei* DMC1 is most closely related to the Rad51 homologue Dmc1, which has been shown in a number of organisms to have a meiosis-specific function (Bishop *et al.*, 1992; Ding *et al.*, 2001; Doutriaux *et al.*, 1998; Fukushima *et al.*, 2000; Gupta *et al.*, 2001). In yeast, mutation of *dmc1* has been shown to result in a defect in reciprocal recombination and synaptonemal complex formation during meiosis suggesting that it is involved in DNA strand exchange (Bishop *et al.*, 1992). The observations that *dmc1* mutants have reduced levels of interchromosomal recombination and increased levels of intrachromosomal recombination (Bishop *et al.*, 1992), and that Rad51 foci form in *dmc1* mutants, whereas Dmc1 foci fail to form in *rad51* mutants (Bishop, 1994), suggests that the role of Dmc1 may be to direct recombination to occur between homologous chromosomes and not sister chromatids. In mice, *dmc1* knockout mutants are viable but sterile, with reproductive organs smaller than normal (Masson and West, 2001). Much of the biochemical analysis of the function of Dmc1 has been carried out using purified recombinant human Dmc1 (Masson and West, 2001). Human Dmc1 has been shown to bind single strand DNA and double stranded DNA as a linear array of stacked octomeric rings (Masson *et al.*, 1999; Passy *et al.*, 1999), and to exhibit weak strand transfer activity, with a 5′ - 3′ preference (Masson *et al.*, 1999).

Surprisingly, the expression of Dmc1 is not always limited to cells undergoing meiosis. Two rice Dmc1 homologues have been identified that appear to be differentially expressed during meiosis and mitosis, suggesting that their function may not be limited to meiotic recombination (Kathiresan *et al.*, 2002). Despite this observation, no mitotic role for rice Dmc1 has been described to date. In fact, in most organisms other than rice, including humans, mice and yeast (Fukushima *et al.*, 2000; Gupta *et al.*, 2001; Yoshida *et al.*, 1998), the expression of Dmc1 is meiosis specific.

T. brucei has been shown to be able to undergo genetic exchange. The first evidence for this was the production of hybrid trypanosome strains by the co-transmission of two parental strains through tsetse flies (Jenni *et al.*, 1986). Observations of the segregation of genetic markers in these transmissions, the increased frequency of chromosomal recombination in hybrids relative to parental strains, and the generation of hybrids with either 2N or 3N DNA content, but not intermediate amounts, all suggest that meiotic division is involved, despite the fact that a cell stage with haploid DNA content has not been detected (Gibson, 2001). Our analysis of all the *RAD51*-like genes was carried out in monomorphic bloodstream form cells that cannot be transmitted through tsetse flies and enter the life-cycle stage where genetic exchange is proposed to take place (Gibson and Whittington, 1993; Tait *et al.*, 1989). If *T. brucei* DMC1 is indeed meiosis-specific, the absence of any discernable phenotypes for *DMC1* mutants is not surprising.

One difficulty with this interpretation is that we can detect DMC1 mRNA (Section 3.8.1). The simplest explanation for these data is that, although *DMC1* is transcribed, the regulation of its expression is post-transcriptional. To date, we have not tested if *DMC1* is expressed at the protein level, because of the lack of any antibody directed against the protein. Western or *in situ* hybridisation analysis would allow us to determine whether or not DMC1 is expressed in Lister 427 bloodstream form cells. If *DMC1* does function in *T. brucei* genetic exchange, it might be expected that it would be expressed at the protein level only in epimastigote cells. This could be tested by Western analysis of whole cell extracts of transgenic epimastigote cells containing an epitope-tagged copy of *DMC1*. Alternatively, if a DMC1 antibody was available, Western analysis of wild type epimastigote cells could be conducted. As it has been shown in yeast, Dmc1 forms foci during meiosis (Shinohara *et al.*, 2000), as Rad51 does in mitotic and meiotic cells (Lisby *et al.*, 2004; Shinohara *et al.*, 2000), *in situ* hybridisation could also be carried out. This would avoid potential problems of large numbers of epimastigote cells required to run Western analysis. If DMC1 does act in meiosis, another expectation would be that *T. brucei* DMC1 and RAD51 foci should be

observed to co-localise, as this has been seen in other organisms (Masson *et al.*, 1999; Masson and West, 2001; Shinohara *et al.*, 2000) and the two proteins appear to act in strand exchange. To more definitively test that DMC1 has a meiosis-specific function is more complex. This would require disruption of *DMC1* in two fly-transmissible strains of *T. brucei* and their co-transmission through flies. If DMC1 is involved in genetic exchange, mini-satellite analysis of the progeny (MacLeod *et al.*, 1999) could theoretically be used to determine if the mutation of *DMC1* caused a reduction in the ratio of recombinant genotypes relative to the parental genotypes, compared to the ratio obtained after the co-transmission of two *DMC1* wild type strains. The co-transmission of two parental strains is required so that genetic exchange results in a genotypic change. This work has a number of problems. Transmission of a single parental stock has a very low efficiency and, as a result, the successful co-transmission of two parental strains is likely to be even less efficient although it has been observed (MacLeod *et al.*, 1999). Another difficulty is the generation of the *DMC1* mutant cell lines in fly-transmissible *T. brucei* cells. The generation of *DMC1* mutants in Lister 427 bloodstream for cells, regarded as the easiest strain to work with, was much more laborious than first thought. Therefore, the generation of mutants in fly-transmissible strains, which are more difficult to work with, may prove to be an arduous process.

6.3 *RAD51-3* and *RAD51-5*

6.3.1 The roles of *RAD51-3* and *RAD51-5* in DNA repair and recombination

The genetic analysis in this project has identified two further factors, beyond RAD51, that are central to *T. brucei* DNA repair and recombination: RAD51-3 and RAD51-5. The DNA damage sensitivities observed in *rad51-3* -/- and *rad51-5* -/- 427 bloodstream stage mutant cells, relative to their heterozygote antecedents and wild type cells, were similar to each other and possibly to that of *rad51* -/- mutants (McCulloch and Barry, 1999). In contrast, the observed defects in homologous recombination in *rad51-3* -/- and *rad51-5* -/- cells appeared not to be as severe as those observed in *rad51* -/- cells. These results may indicate that the roles of *RAD51-3* and *RAD51-5* are not as essential as that of *RAD51* in DNA repair and recombination. However, comparisons between the DNA damage sensitivities observed in *rad51-3* -/- and *rad51-5* -/- mutant cells compared with those observed in *rad51* -/- cells is not straightforward, as the DNA damage assays

carried out were slightly different. Those carried out for *RAD51-3* and *RAD51-5* were DNA damage survival assays, following the ability of single cells to grow out in the presence of MMS. In contrast, those carried out for *RAD51* were simply growth curves in the presence of MMS (McCulloch and Barry, 1999). Similarly, the DNA transformation assays to measure the recombination efficiencies of the *RAD51-3* and *RAD51-5* mutants are difficult to compare with the *RAD51* analysis, as different researchers have conducted them at different times. The question of the extent of the severity of DNA damage repair and recombination defects in the *RAD51-3* and *RAD51-5* mutants is important as it directs us towards understanding their potential functions. If RAD51-3 and RAD51-5 are not as essential as that of RAD51 in these processes it may suggest that they act as co-factors of RAD51. For instance, Rad55 and Rad57 mutants in *S. cerevisiae* have mild repair and recombination defects relative to Rad51 (Symington, 2002). Moreover, over-expression of Rad51 can alleviate these defects (Hays *et al.*, 1995) and in fact, some mutations of Rad51 can partially by-pass the requirement for Rad55 and Rad57 in DNA repair (Fortin and Symington, 2002). In a similar manner, mutation of any of the five vertebrate *Rad51*-like genes (*Rad51B*, *Rad51C*, *Rad51D*, *Xrcc2* and *Xrcc3*) in chicken DT40 cells appears to create less severe repair and recombination defects than mutation of *Rad51*: whereas the *Rad51*-like mutants are viable, *Rad51* mutants are lethal (Takata *et al.*, 2001), and, expression of human *Rad51* in the DT40 *Rad51*-like mutant appeared to suppress the phenotypes. However, whereas there is direct biochemical evidence for the *S. cerevisiae* Rad55/57 heterodimer aiding in the process of Rad51 nucleoprotein filament formation and strand exchange (Johnson and Symington, 1995; Sung, 1997b), the same direct involvement of the vertebrate Rad51-like proteins has not been determined.

As an alternative to *T. brucei* RAD51-3 or RAD51-5 functioning as RAD51 co-factors, we can consider whether or not the proteins might operate as RAD51-independent enzymes in a different pathway of recombination. We already know that at least one back-up pathway of recombination operates in the absence of RAD51 in *T. brucei* (Conway *et al.*, 2002b), and acts upon short stretches of imperfect sequence homology (Conway *et al.*, 2002b). However, this appears to be a relatively minor pathway as it has never been seen in RAD51-proficent *T. brucei* cells, and the frequency of recombination in the absence of RAD51, when this pathway is revealed, is around 10-fold less than the wild type cells. The severity of the impairment of recombination in the *RAD51-3* and *RAD51-5* mutants (4 – 9 fold; Section 4.6) seems incompatible with either of them directing microhomology-dependent recombination. Nevertheless, it is interesting to

note that in *RAD51*, *RAD51-3* or *RAD51-5* mutants the observed patterns of most plasmid integrations are consistent with homologous recombination (see Section 4.6; (Conway *et al.*, 2002b). We cannot, therefore, exclude that a RAD51-independent pathway of recombination, reliant on longer stretches of homology, does indeed exist in *T. brucei*. Addressing this question would require the generation of double mutants of *RAD51*, *RAD51-3* and *RAD51-5*, and perhaps the other *RAD51*-like genes not analysed in this project. The best evidence we currently have that RAD51-3 and RAD51-5 function in aiding *T. brucei* RAD51 comes from *in situ* hybridisation analysis (see below).

The DNA repair and recombination phenotypes observed in *rad51-3 -/-* and *rad51-5 -/-* mutant cells would appear to suggest that they are more central to general repair and recombination in *T. brucei* than either MRE11 or KU70/80. *mre11 -/-* cell lines were found to have, at best, a slightly increased DNA damage sensitivity, and a 3-4 fold reduction in recombination relative to the wild type and *MRE11 +/-* cells (Robinson *et al.*, 2002; Tan *et al.*, 2002). *T. brucei ku70 -/-* or *ku80 -/-* mutants display no detectable DNA damage sensitivity or recombination defect (Conway *et al.*, 2002a; Janzen *et al.*, 2004). These findings are in keeping with the suggestion that homologous recombination, including the functions of RAD51-3 and RAD51-5, is the main route of DNA repair in *T. brucei*. Whether or not *rad51-3 -/-* and *rad51-5 -/-* mutant cells show gross chromosomal rearrangements or telomere defects, as has been described for the *mre11 -/-* and *ku70 -/-* mutants respectively, remains to be determined. However, it has been shown that mutation of RAD51-like genes in chicken DT40 cells causes chromosome abnormalities (Takata *et al.*, 2001), and it has been suggested that human RAD51D has a role in telomere protection (Tarsounas *et al.*, 2004b).

Given the very similar DNA repair and recombination defects in *RAD51-3* and *RAD51-5* mutants, an important question is do the two proteins operate in the same or distinct pathways, and do they act with each other or have separate roles? In order to attempt to answer this question, RAD51 *in situ* hybridisation analysis in response to DNA damage was carried out in the *RAD51-3* and *RAD51-5* mutants. The results of this analysis appeared to show that both contribute to the reorganisation of RAD51 sub-cellular localisation into discrete RAD51 foci following phleomycin-induced DNA damage. A reduced percentage of cells displaying *RAD51* foci, and a reduced average number of foci per cell, was observed in both *rad51-3 -/-* and *rad51-5 -/-* mutant cells. This suggests that both proteins contribute to an evolutionarily conserved pathway of RAD51 relocalisation in response to DNA damage, and that these RAD51 foci represent DNA

repair centres that have been described in yeast and mammals, in both mitotic and meiotic cells (Gasior *et al.*, 1998; Haaf *et al.*, 1995; Haaf *et al.*, 1999; Lisby *et al.*, 2001; Lisby *et al.*, 2003a; Lisby *et al.*, 2003b; Lisby *et al.*, 2004; Lisby and Rothstein, 2004; Lukas *et al.*, 2003; Maser *et al.*, 1997; Melo *et al.*, 2001; Raderschall *et al.*, 1999). We have not, however, determined whether the foci are centres of multiple DNA breaks, as has been suggested for other organisms (Lisby *et al.*, 2001).

RAD51 foci could still be observed in both the *rad51-3* -/- and *rad51-5* -/- cell lines. This could mean one of two things. First, the 'residual' foci could arise in response to spontaneous, rather than induced, DNA damage, and this may not require the actions of RAD51 co-factors. Spontaneous DNA damage results from the natural metabolism of the cells, which can result in the production of reactive oxygen species or DSBs created during DNA replication and transcription (Kupiec, 2000). Induced damage, on the other hand, is caused by an external source, such as, UV irradiation, MMS or phleomycin. It been shown in mammals that the formation of DNA damage-induced Rad51 foci requires the actions of Brca2 and the Rad51 paralogues, whereas the formation of spontaneous foci has no such requirement (Tarsounas *et al.*, 2003). In yeast, it has been shown that Rad51 foci do not form in *rad52* mutant cells, whereas mutation of *rad54*, *rad55* and *rad59* has no effect on the formation of Rad51 foci (Lisby *et al.*, 2004). This is direct contrast to vertebrate cells, where mutation of any of the Rad51-paralogues results in a failure to form Rad51 foci (Tarsounas *et al.*, 2003). In addition, Rad52 in yeast has been shown to form foci in response to induced and spontaneous DNA damage (Lisby *et al.*, 2001). An alternative possibility for the residual RAD51 foci in *T. brucei RAD51-3* and *RAD51-5* mutants is that that the DNA damage response system in *T. brucei* is degenerate, and RAD51-3 and RAD51-5 contribute to distinct pathway, which lead to the formation of repair centres.

From the *in situ* hybridisation results, we can conclude only that RAD51-3 and RAD51-5 operate in broadly the same capacity in aiding RAD51 foci formation. We cannot make any predictions regarding their biochemical functions, nor whether those functions are the same or distinct. For example, it is conceivable that one of the *T. brucei* proteins might be a signalling protein, to identify DNA damage, whilst the other directly interacts with RAD51. This could be addressed by testing their interactions with RAD51 using either yeast two-hybrid analysis or by co-immunoprecipitation experiments. RAD51-3 and RAD51-5 might also be involved in separate repair complexes, with potentially over-lapping functions. Such a scenario would place the *T. brucei* RAD51-like protein complement in the context of mammalian cells, where at

least two complexes of Rad51 paralogues have been identified (Liu *et al.*, 2002; Masson *et al.*, 2001a; Masson *et al.*, 2001b; Wiese *et al.*, 2002; Fig. 1.9), rather than the single Rad55/57 heterodimer found in yeast (Symington, 2002). Testing of this prediction could be achieved by the generation of *RAD51-3* and *RAD51-5* double mutants, or testing interactions amongst the RAD51-like proteins using yeast two-hybrid or co-immunoprecipitation analyses. Finally, RAD51-3 and RAD51-5 might operate at entirely distinct steps in the pathway of RAD51 foci formation, and indeed in DSB repair. Biochemical characterisation of the purified enzymes activities would most rigorously define whether or not this was the case. The observation that *T. brucei rad51-3 -/-* mutant cells show a reduced frequency of VSG switching, whereas *rad51-5 -/-* mutant cells do not, may lend credence to this possibility (see below).

6.3.2 The roles of *RAD51-3* and *RAD51-5* in antigenic variation

Perhaps the most surprising observation from this work is that, despite *RAD51-3* and *RAD51-5* mutants having very similar DNA repair and recombination phenotypes, only *RAD51-3* mutants display a defect in antigenic variation. The most convenient explanation for this result is that VSG switching is a highly variable process, and the *in vivo* assay used to analyse the switching frequency has failed to detect a role for RAD51-5, or that RAD51-5 has a less important function than RAD51 or RAD51-3 in this reaction and the assay is too insensitive to detect a small reduction in VSG switching in *rad51-5 -/-* mutants. It is difficult to address the former possibility, but it seems unlikely that RAD51-5 is less important than RAD51-3 in RAD51 function. In the *in situ* hybridisation analysis, the *rad51-5 -/-* mutants appeared, if anything, to have a more severe defect than the *rad51-3-/-* mutants during RAD51 foci formation (Section 5.2.2). It is also unlikely that the *RAD51-3* mutants have a more severe growth defect that could serve to artificially identify a VSG switching defect, since the *RAD51-5* mutants showed at least as significant a level of growth impairment in the *in vivo* growth assay as the *RAD51-3* mutants (Section 4.3.2).

An alternative explanation for these data is that *T. brucei* antigenic variation is carried out by a specific or unusual form of homologous recombination that utilises only a discrete subset of proteins from the general DNA repair and recombination machinery. Perhaps *T. brucei* has evolved a specialised homologous recombination pathway for the process of antigenic variation to avoid alterations to the genome during the movement

of *VSGs*. This suggestion of a specialised homologous recombination machinery is lent weight by the observation that MRE11, like RAD51-5, has no detectable function in antigenic variation, despite contributing to DNA repair and recombination (Robinson *et al.*, 2002). Similarly, although mismatch repair is known to constrain *T. brucei* homologous recombination to occur only between sufficiently related DNA sequences, it appears not to constrain antigenic variation (Bell *et al.*, 2004). If correct, this hypothesis would suggest distinct roles for RAD51-3 and RAD51-5 in RAD51-catalysed strand exchange. Recent analysis of the functions of the *Rad51*-like genes in vertebrates is compatible with this suggestion. As has been described previously, the Rad51-like proteins in vertebrates appear to exist in two, and perhaps three, sub-cellular complexes (Liu *et al.*, 2002; Masson *et al.*, 2001a; Masson *et al.*, 2001b; Wiese *et al.*, 2002). There are now suggestions that these may have functions beyond simply directly loading the RAD51 nucleoprotein filament onto single strand DNA, as the yeast Rad55/57 heterodimer is thought to do (Sung, 1997b). Xrcc3 in mammals has been shown to act very early in the pathway of DSB repair, based on the findings that Xrcc3 re-localises to DNA DSBs within 10 minutes, independently of Rad51 (Forget *et al.*, 2004). The fact that Rad51C forms a complex with Xrcc3 (Masson *et al.*, 2001a) would make it another suitable candidate to have an early role in DSB repair. Rad51C has been shown to be closely associated with a Holliday junction resolution and branch migration activity identified in mammalian cell extracts (Liu *et al.*, 2004), suggesting that it can act at late steps of homologous recombination, following Rad51-catalysed strand exchange. Finally, Rad51D is thought to act at telomeres, and whether this is related to Rad51 functioning is unclear (Tarsounas *et al.*, 2004b).

If the regulation of homologous recombination in *T. brucei* operates in a manner more akin to mammalian recombination, with the RAD51-like proteins having functions beyond those of simply co-factors, then it is conceivable that RAD51-3 and RAD51-5 will have non-equivalent functions in antigenic variation. For example, RAD51-5 may function in either late or early steps in homologous recombination that are by-passed during the process of antigenic variation. Perhaps *T. brucei* RAD51-5, like human Rad51C, is needed mainly for Holliday junction processing, and this intermediate is not used during VSG switching, as has been suggested by two models for SDSA (Barry, 1997; Borst *et al.*, 1996). Unfortunately, this suggestion seems incompatible with the role of RAD51-5 in promoting *T. brucei* RAD51 foci formation. Alternatively, RAD51-5 may act early to prepare DNA breaks for RAD51 binding, and specific breaks initiate VSG switching that by-pass this function (see below). In contrast to this, RAD51-3 may

function, as it Rad55/57 does in yeast, to directly aid the formation of *T. brucei* RAD51 nucleoprotein filaments, a process retained by the antigenic variation recombination machinery, and therefore explains the reduction in VSG switching observed in *rad51-3* -/- mutants cells.

Currently we know essentially nothing about the initiation of *T. brucei* VSG switching. The high rate of switching in pleomorphic *T. brucei* cells has prompted Barry (1997) to suggest that a 70-bp repeat-specific endonuclease must create breaks which direct VSG recombination. Alternatively, Borst (1996) suggested that VSG switching relies on random DNA breaks. It is clear from a number of organisms that specific enzymes to catalyse DNA breaks can initiate directed forms of homologous recombination. For example, meiotic recombination, driven by Rad51 and Dmc1, is directed to specific sites of DSBs created by the enzyme Spo11 (Keeney *et al.*, 1997; Romanienko and Camerini-Otero, 2000). Indeed, two Spo11 homologues have been identified in the *T. brucei* genome (R. McCulloch, pers. comm.); the possibility that this duplication has created one enzyme for VSG switching and one for meiosis seems remote, however, since *T. cruzi* and *L. major* both contain two Spo11 homologues also. In *S. cerevisiae* mating type switching, homologous recombination is directed to the MAT locus by cleavage *via* the HO endonuclease (Haber, 2002). In both cases, the DSBs appear to be processed by Mre11/Rad50/Xrs2(Nbs1) to facilitate recombination. However, initiation by enzyme directed DSBs is not universal. In *S. pombe*, mating type switching is initiated by a combination of replication fork pausing and an unidentified single-strand lesion (Dalgaard and Klar, 2001; Vengrova and Dalgaard, 2004). The genetic components that contribute to *S. pombe* mating type switching have not been determined. It remains possible, therefore, that VSG switching could have an initiating event that by-passes the function of MRE11 in DSB processing, and potentially RAD51-5.

Other, more speculative, possibilities for the non-involvement of RAD51-5 or MRE11 in catalysing VSG switching can be considered. The ESB may have evolved to regulate or limit the number of recombination proteins that gain access to the active BES (Navarro and Gull, 2001), or that a specialised antigenic variation homologous recombination machinery evolved because of this unusual structure. The ESB may perform the role of global regulator of VSG switching, only allowing particular recombination reactions to occur within it, and perhaps co-ordinating VSG switches with DNA replication. It is also possible that base J present within the inactive BESs (van Leeuwen *et al.*, 1997) plays a role in VSG switching, although it is hard to predict

how a modified base would affect VSG switching or regulate homologous recombination.

A more extreme hypothesis to explain the complex picture of factors which act, and do not act, in antigenic variation may be that VSG recombination during switching is not directly catalysed by RAD51. It may be that RAD51 acts purely in a regulatory, or signalling, capacity and controls a wholly novel, as yet uncharacterised, recombination machinery. If this were the case, then presumably RAD51-3 acts in conjunction with RAD51 in such a role, which has no requirement for RAD51-5. However, a role of this nature for RAD51 is unprecedented, making it an unlikely possibility. Indeed, the available evidence suggests that Mre11, as well as a number of repair kinases, act to signal damage and co-ordinate DSB repair, with Rad51 functioning purely in a catalytic function (Lisby *et al.*, 2004).

6.4 *RAD51-3* and a role for homologous recombination in *in situ* switching?

In the analysis of VSG switching in *rad51* -/- cells, McCulloch and Barry (1999) identified a perplexing result: the reduced frequency of antigenic variation appeared not only to be caused by reduction in *VSG* recombination, but also by reduced transcriptional, or *in situ*, switching. The results for the *RAD51-3* mutant analysis has identified an identical phenotype, since the ratio of VSG or ES gene conversions relative to *in situ* switches is not changed, despite the reduction in VSG switching frequency (Section 4.8.1). On the face of it, these data makes the involvement of recombinational proteins in *in situ* switching a more likely possibility.

So, is *in situ* switching actually a form of homologous recombination catalysed by RAD51, and perhaps RAD51-3? The answer to this is almost certainly 'no'. Previous analyses of *in situ* switches, despite occasionally finding DNA rearrangements, have seen no consistent DNA rearrangements (Horn and Cross, 1997) and in some cases no detectable rearrangements (Navarro and Cross, 1996). Furthermore, the VSG promoter has been changed to the rDNA promoter (Rudenko *et al.*, 1995) and antibiotic makers have been inserted around the promoter (Chaves *et al.*, 1999; Navarro and Cross, 1996), without these sequences changing during switching, therefore ruling out gene conversions.

An alternative explanation for the *RAD51* and *RAD51-3* mutant phenotypes may be that *in situ* switching occurs by a homology sensing mechanism, perhaps functioning to

facilitate the movement of a silent BES into, and the active BES out of, the ESB. Very little is known about the function of the ESB (Chaves *et al.*, 1998; Navarro and Gull, 2001), and to date we cannot even say whether it has a direct role in regulating VSG switching, or is simply the site of BES transcription. The possibility that RAD51 and RAD51-3, and hence homologous interactions, contribute to in *situ* switching is perhaps supported by the finding that the process appear to be a co-ordinated, rather than a spontaneous, event. The evidence for this is two-fold. First, the insertion of two antibiotic markers in different BESs allowed for the selection of *in situ* switches, and the isolation of an unstable double resistant phenotype, in which transcription rapidly switched between the two BESs (Chaves *et al.*, 1999). Second, using three marked BESs, it was determined that putative switch intermediates involving three BESs are never recovered (Ulbert *et al.*, 2002a). These results suggest that circumstances in which two BESs are transcribed probably represents a natural intermediate that occurs during *in situ* switching. Furthermore, the process appears to be coordinated, as only two BESs are ever involved, the one being inactivated and the one being activated (Chaves *et al.*, 1999; Ulbert *et al.*, 2002a). Direct analysis of the potential roles of RAD51 and RAD51-3 in *in situ* switching should be feasible. For instance, using a strain with two BESs marked with antibiotic genes to select for *in situ* switching events (Chaves *et al.*, 1999), and generating *RAD51* or *RAD51-3* mutants in that strain, would allow us to genetically define the roles of RAD51 and RAD51-3. Co-localisation of RAD51 or RAD51-3 with a GFP-tagged BES (Navarro and Gull, 2001) could assess the likelihood of RAD51 and RAD51-3 associating with the ESB. Finally, a role for RAD51 or RAD51-3 in *in situ* switching could be defined using the *T. brucei* strain that allows for isolating putatively switching intermediates between two BESs (Ulbert *et al.*, 2002b) if mutants fail to form intermediates.

6.5 The DNA damage response in *T. brucei*

The analysis of gene and protein expression following DNA damage in this thesis suggests that there is no increased quantity of *RAD51*, *RAD51-3*, *RAD51-5* or *DMC1* mRNAs following phleomycin treatment, and no up-regulation of RAD51 at the protein level. This is an unusual result as most micro-organisms, including *L. major* (McKean *et al.*, 2001) and *P. falciparum* (Bhattacharyya and Kumar, 2003), induce *RAD51* expression in response to DNA damage. Many different DNA damaging agents can

result in the increased expression of the proteins involved in DNA repair, including ionising and UV-irradiation, and chemicals such as MMS, mitomycin C, phleomycin and MNNG (Kupiec, 2000). The increase in expression resulting from exposure to one of these agents can occur at the RNA or the protein level. Rad51 RNA and protein levels increase 3-6 fold in *S. cerevisiae* following exposure to γ-rays and MMS (Basile *et al.*, 1992; Shinohara *et al.*, 1992), whilst, in *S. pombe* a 3 – 5 fold increase in RNA level has been described (Jang *et al.*, 1994). In response to UV and MMS *E. coli* and *Tetrahymena thermophila* also show increases in RNA and protein levels, 15-fold and 30 – 100 fold respectively (Campbell and Romero, 1998; Little and Mount, 1982; Walker, 1984).

Up-regulation of Rad51 is not universal, however, as no evidence for this has been found in mammals (Tarsounas *et al.*, 2004a), some archae (Komori *et al.*, 2000) and *Neisseria gonorrhoeae* (Black *et al.*, 1998). The explanations for this appear to be variable. Mammals appear to favour NHEJ over homologous recombination, at least in G1 cells, to repair DNA damage, probably accounting for the lack of increased Rad51 expression in response to DNA damage. The archae that do not increase expression of RadA in response to DNA damage are thermophilles and are therefore probably constantly exposed to DNA damaging conditions and requires a consistently high level of expression of RadA. The lack of regulation of RecA in *Neisseria* possibly results from a switch in the way that DNA damage is repaired, a suggestion based on the observations that the RecA and NER proteins do not respond to DNA damage (Black *et al.*, 1998).

So, why does *T. brucei* appear not to regulate *RAD51* expression in response to DNA damage when a highly related kinetoplastid, *L. major*, does so? It seems unlikely that *T. brucei* is subject to high levels of environmental DNA damage, and a more likely explanation lies in antigenic variation. If antigenic variation utilises RAD51 in the catalysis of VSG switching, which can occur at high rates in bloodstream stage cells (Turner and Barry, 1989), then RAD51 protein levels may need to be maintained at higher levels that if it were only required for DSB repair. The induction of VSG switching may be seen as a form of continuous DNA damage and, hence, *T. brucei* has lost the mechanism to respond to induced damage by increased expression of RAD51, such as occurs in *L. major* and *P. falciparum*. In performing the *in situ* hybridisation analysis of RAD51 localisation an interesting finding emerged which may be related to this. It appears that bloodstream stage cells appeared able to continue their progression through the cell cycle, despite being exposed to DNA damage. In other organisms,

exposure to DNA damage agents causes an arrest at the G2/M boundary (Paques and Haber, 1999). Richard McCulloch carried out DAPI analysis on bloodstream stage cells exposed to increasing concentrations of phleomycin, and found no evidence for an alteration of the ratio of parasite cell types in different life cycle stages (Fig. 6.1A). Interestingly, when he carried out the same assay using procyclic cells, the pattern of cell cycle forms did alter (Fig 6.1B), with a reduction in cells undergoing division (2 nuclei and 2 kinetoplast; 2N 2K), a reduction in G1 phase cells (1N 1K), a large increase in abnormal cells and a small increase in cells that had completed kinetoplast division (1N 2K). This seems compatible with a transient block at the G2/M phase boundary, something not seen in bloodstream form cells.

The explanation for the above data could either be that bloodstream stage cells arrest at any stage of the cell cycle, or traverse through cell division with un-repaired DNA damage. In an attempt to address this question, microscopic images of wild type bloodstream form cells exposed to phleomycin were searched for cells that were in the process of nuclear division (Fig. 2). RAD51 foci could be readily detected in all life cycle stages, including examples with dividing nuclei, and the foci appeared to be segregated. Clearly, this result depicts fixed cells, and imaging of RAD51 in live, dividing cells would be required to confirm it. However, it suggests the possibility that bloodstream stage cells generate DSBs during VSG switching, and do not arrest as this could result in immune destruction. Furthermore, it would infer that bloodstream stage cells cannot distinguish between generalised DNA damage and that needed for VSG switching, implying that VSG switching would be a component of general homologous recombination, and not a specialised RAD51-independent reaction. Perhaps also they have an undefined checkpoint response to DNA damage removed or down-regulated, and are therefore predisposed to continue the cell cycle.

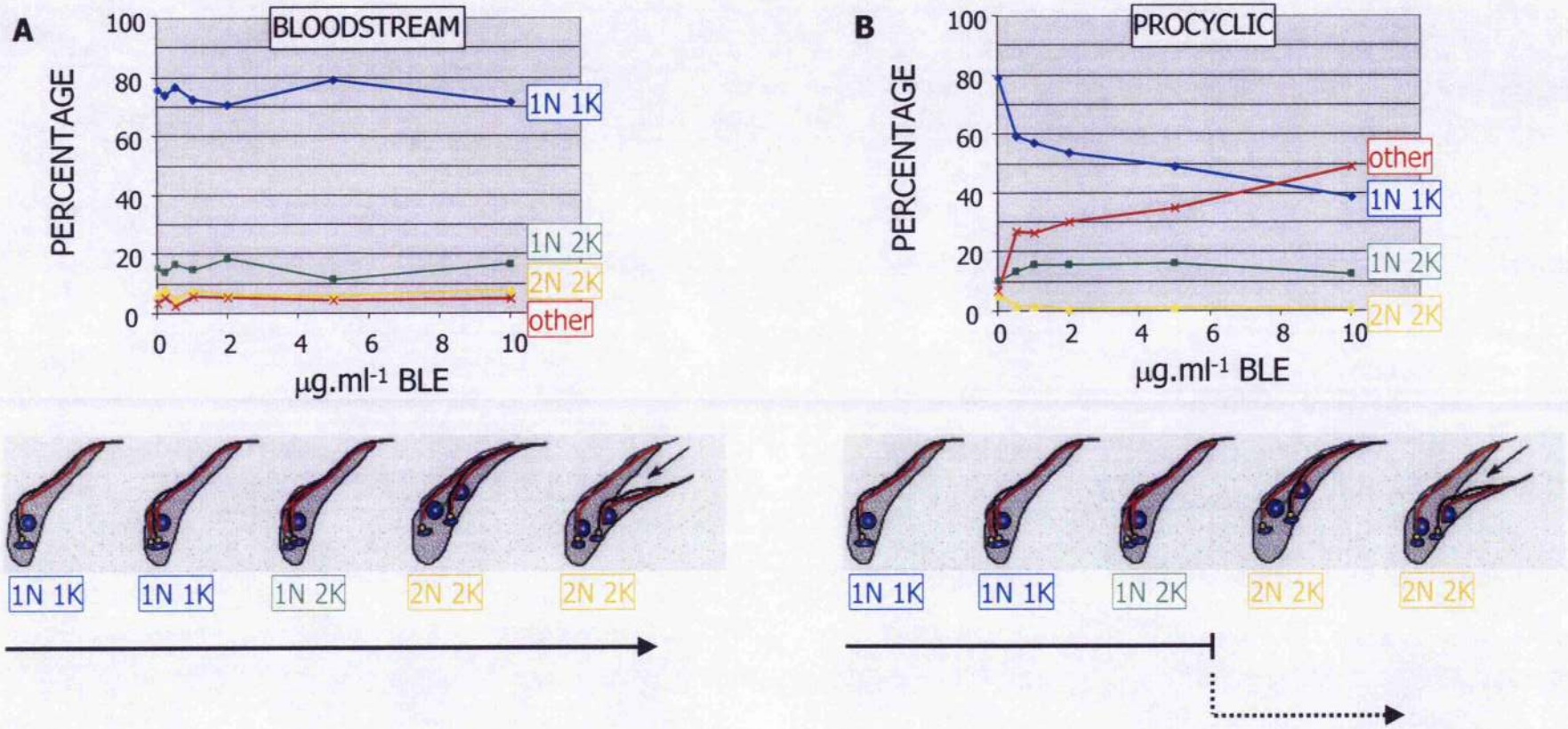


Figure 6.1: The effects of DNA damage on the cell cycle of 427 bloodstream form and procyclic cells. Results are shown for DAPI staining carried out after exposure to Phleomycin for 18 hours for (A) bloodstream form cells and (B) procyclic cells. The various cell cycle stages are depicted below the graphs, the black arrows represent the interpreted effect of DNA damage on the cell cycle. 1N 1K: 1 nucleus and 1 kinetoplast. 1N 2K: 1 nucleus and 2 kinetoplast. 2N 2K: 2 nuclei and 2 kinetoplast. other: cells that have a DNA content that does not fit one of the previous categories. Figure adapted from Mckean (2003)

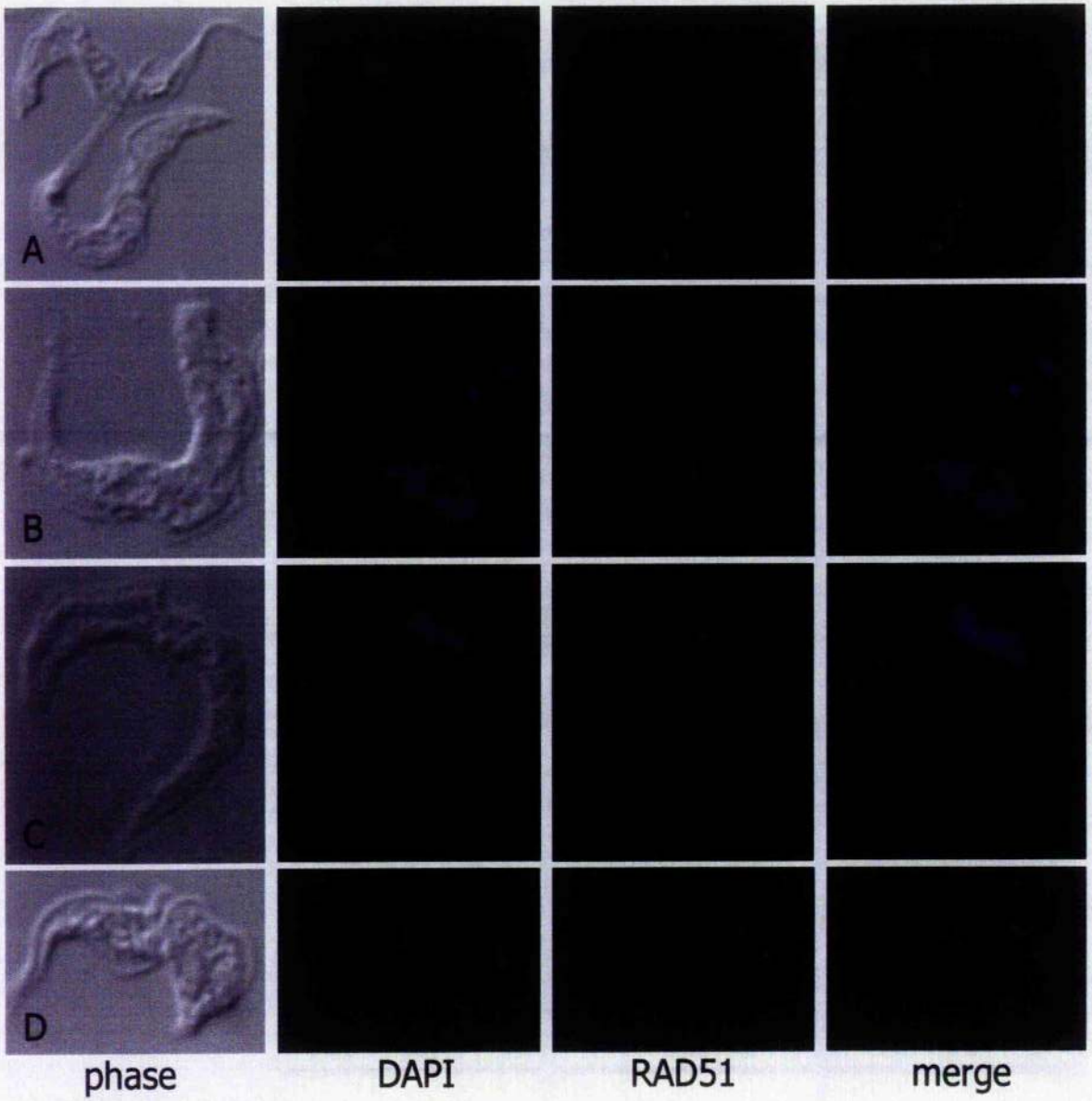


Figure 6.2: RAD51 foci formation in wild type cells in the process of nuclear division. Each cell or group of cells is shown in phase contrast (phase), after staining with DAPI (DAPI) or after hybridisation with anti-RAD51 antiserum and secondary hybridisation with FITC conjugate (RAD51). Merged images of DAPI and FITC stained cells are shown also. All cells were exposed to 1.5 $\mu g.ml^{-1}$ phleomycin. A: two cells one of which is 1N 2K with 2 foci, B: a 2N 2K cell with 2 foci, one in each nucleus, C: a 1N 2K cell with 2 foci, D: a 2N 2K cell with 2 foci, both in the same nucleus.

6.6 Further experiments

The results presented in this thesis leave a number of questions unanswered. Suggested below are a number of experiments that could be carried out in the future to attempt to start answering those questions. Had more time been available, the over-expression experiments described in Chapter 5 would potentially have helped define the roles of the *RAD51*-like genes, by determining whether any of RAD51-3, RAD51-5 or DMC1 can complement a RAD51 defect and hence catalyse strand exchange. All of the constructs for this analysis were generated, but the reduced recombination efficiency of the *rad51* -/- cells proved too big a hurdle to allow the completion of this experiment. One possible way around the recombination deficiency of *rad51*-/- mutant cells would be to generate the over-expression cell lines in wild type cells and subsequently generate *rad51*-/- mutants.

At the outset of this project, one of the aims was to identify any *RAD51*-like genes present in the *T. brucei* genome. This resulted in the identification of 5 further *RAD51*-like genes, only 3 of which have been functionally analysed during this project. Functional characterisation of the two remaining *RAD51*-like genes must be viewed as a priority. Such characterisation, along the lines described in this thesis, would give a fuller picture of DNA repair, recombination and antigenic variation in *T. brucei*.

As I have argued in this discussion, it is possible that RAD51-3 and RAD51-5 have distinct functions in repair and recombination, and there may be multiple repair complexes, analogous to those in mammalian cells. Yeast 2-hybrid analysis or co-immunoprecipitation, involving all the RAD51-like genes, would be necessary to address this question.

Finally, purification of the *RAD51*-like proteins would be a valuable step to allow biochemical characterisation of their activties. The purified proteins could be used to assay for specific functions in DNA repair and homologous recombination. Assays to examine DNA strand exchange, DNA binding and ATP hydrolysis are well established (Kowalczykowski and Eggleston, 1994), and would lead to a more accurate picture of how the RAD51-like proteins are involved in repair, recombination, and antigenic variation in *T. brucei*.

Appendix 1: A list of the oligonucleotides used for PCR.

Primer Name	Sequence	Restriction Sites	
αβ 3'-*Hpa*I	CCGTTAACCT ATTTTCTTTG ATGAAAGGG	*Hpa*I	
ACT3'-*Sph*I		*Sph*I	
βα 5'-*Hpa*I	CCGTTAACTG GGTCCCATTG TTTGCCTC	*Hpa*I	
DMC1-D1	GTAACGGCTC TCTTTCGC		
DMC1-D2	CGCACGAGCA TCAGATGC		
DMC1-D3	CCTGAGCGGA TTAAGCCC		
DMC1-For	CAGATATCGT ATGCAGCACG TGGGAACG	*Eco*RV	
DMC1-LHF-For	GATACACTAG AATTCACTCT ACGCATGAGG ACG	*Eco*RI	
DMC1-LHF-Rev	GGTAAGCTTG TTAACCTCTG TTGCAGGTAA GCC	*Hind*III	*Hpa*I
DMC1-PCR-KO3'	CGCTGATCCC CCCGTCCTTT GCGCAGCGAC AGTCGCGTCG TTGAGGCATG TATTTTATGG CAGCAACG		
DMC1-PCR-KO5'	GGACAACTCT ACGCATGAGG ACGCGGCGCA CACCATCATG GAGATTGACC TGGGTCCCAT TGTTTGCCTC		
DMC1-Rev	CAGATATCGT ACTCACGTGC GTCAACAATC CC	*Eco*RV	
DMC1-RHF-For	ACGAAGCTTG ATATCATTTT TCTGGCAGAG GGG	*Hind*III	*Eco*RV
DMC1-RHF-Rev	ACTGGCAGAC TCGAGTACAC ACGCGCTGAT CCC	*Xho*I	
DMC1-U1	CGCCTCCGTA ATTGACATGC		
HYGRO-3'	CTATTCCTTT GCCCTCGGAC		
HYGRO-5'	ATGAAAAAGC CTGAACTCAC C		
MCM-3'	GGAATAAAAG GACTCGGG		
MCM-5'	GTGATTTATC ACCACGGC		
NEO-3'	TCAAGAAGGC GATAGAAGGC		
NEO-5'	CGCATGATTG AACAAGATGG		
POLI-3'	CATGCGCCTG TGGTTCAGCA TAGC		
POLI-5'	CAGGAGGATC GTTCGGCACC TTGGC		
P-ORF-3'	TCGAGACAAC TTCAAGGC		
P-ORF-5'	GTCAACAGCA TTATAGCC		
RAD51-For	CAGAATTCGT ATGAACACTC GCACCAAAAA TAAGAAACG	*Eco*RI	
RAD51-Rev	CAGAATTCGT CTAGTCCCTA ACGTCTCCC	*Eco*RI	
RAD51-3-For	CAGATATCGT ATGTCCGTGG AGCAATGC	*Eco*RV	
RAD51-3-KO3'	CCGTGACTTT CAATAACGCC		
RAD51-3-KO5'	TGTAACAACA CTTTGCCG		
RAD51-3-LHF-For	GATACACTAG AATTCCAATG CTCCTCACTT ACCCC	*Eco*RI	
RAD51-3-LHF-Rev	GGTAAGCTTG TTAACAAGAA TATCAAGACT CCGGC	*Hind*III	*Hpa*I
RAD51-3-Rev	CAGATATCGT TCAAAGAGTG GGTCGG	*Eco*RV	
RAD51-3-RHF-For	ACAAAGCTTG ATATCCAAAT GGCACACGTA TGGGG	*Hind*III	*Eco*RV
RAD51-3-RHF-Rev	ACTGGCAGAC TCGAGAGTGG GTCGGAAAAC ATCGC	*Xho*I	
RAD51-5-For	CAGAATTCGT ATGTCTGTGT GTCCTCC	*Eco*RI	
RAD51-5-LHF-For	GATACACTAG AATTCATGTC TGTGTGTCCT CCACC	*Eco*RI	
RAD51-5-LHF-Rev	GGTAAGCTTG TTAACAACAC CTCCAGGACT GTCCC	*Hind*III	*Hpa*I
RAD51-5-Rev	CAGAATTCGT TCAGGGTAAA AAGATGTTTC CC	*Eco*RI	
RAD51-5-RHF Rev	ACTGGCAGAC TCGAGATACG TGTTTTGCAG CCCCG	*Xho*I	
RAD51-5-RHF-For	ACAAAGCTTG ATATCTCTGT TGGTGGTAGT GTGGC	*Hind*III	*Eco*RV
RAD51-5-U2	GGGATTGAGA GAGGACGGG		
TT-LIG-3'	AACCTTTTGG TAAGGGCC		
TT-LIG-5'	ATTCATCGAG TGAGAGCC		
VSG221-3'	TGTATCGGCG ACAACTGCAG		
VSG221-5'	ATGCCTTCCA ATCAGGAGGC		

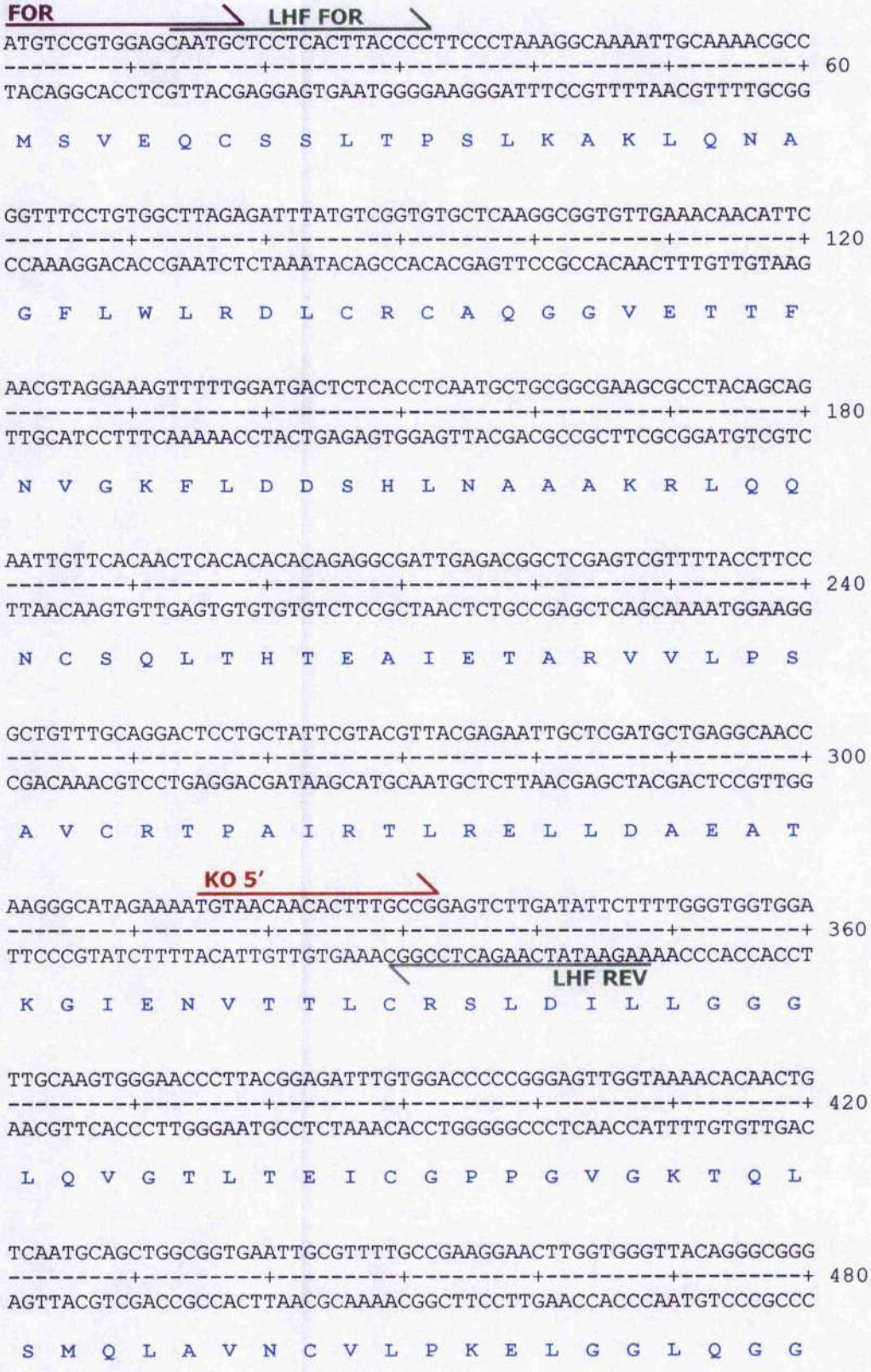
Appendix 2: Accession numbers.

T.b	RAD51	AF174136
	RAD51-3	Tb11.02.0150
	RAD51-4	Tb11.02.4880
	RAD51-5	Tb10.389.1770
	RAD51-6	Tb03.5L5.340
	DMC1	Tb09.211.1210
L.m	RAD51	O61127
	RAD51-3	LmjF33.2490
	RAD51-4	LmjF11.0230
	RAD51-6	LmjF29.0450
	DMC1	O61128
T.c	RAD51	Tc00.1047053508817.50 *
		Tc00.1047053503801.30
	RAD51-3	Tc00.1047053504153.220
	RAD51-4	Tc00.1047053503613.30 *
		Tc00.1047053511165.60
	RAD51-5	Tc00.1047053511837.50 *
		Tc00.1047053510123.30
	RAD51-6	Tc00.1047053508075.20 *
		Tc00.1047053509643.80
	DMC1	Tc00.1047053510729.110 *
		Tc00.1047053506885.310
P.f	RAD51	AF452489
	DMC1	AF356553
E.c	RECA	P03017
C.j	RECA	P42440
N.g	RECA	P21152
B.s	RECA	P16971
S.l	RECA	P48294
S.a	RECA	Q02350
U.m	RAD51	UM03290
	REC2	UM03095
H.s	RAD51	Q06609
	RAD51C	O43503
	RAD51L1	O15315
	RAD51L3	O75771
	XRCC2	O43543
	XRCC3	O43542
	DMC1	Q14565
S.c	RAD51	X64270
	RAD55	Z46796
	RAD57	M65061
	DMC1	M87549
S.p	RPH51	Z22691
	RPH55	AF053410
	RPH57	AB024744
	DMC1	AB008545
A.t	RAD51	U43528
	RAD51Cα	AB062456 *
	RAD51Cβ	AB073493
	XRCC3α	AB062455 *
	XRCC3β	AB073492
	DMC1	U76670
	mtRECA	AY072877
	ctRECA	M98039
O.s	DMC1A	AF265548
	DMC1B	AF265549
D.m	RAD51	BAA04580
	RAD51C	NP_610466
	RAD51D	NP_5773302
	SPNB	NP_476740
	SPND	FBgn0003482 ∞

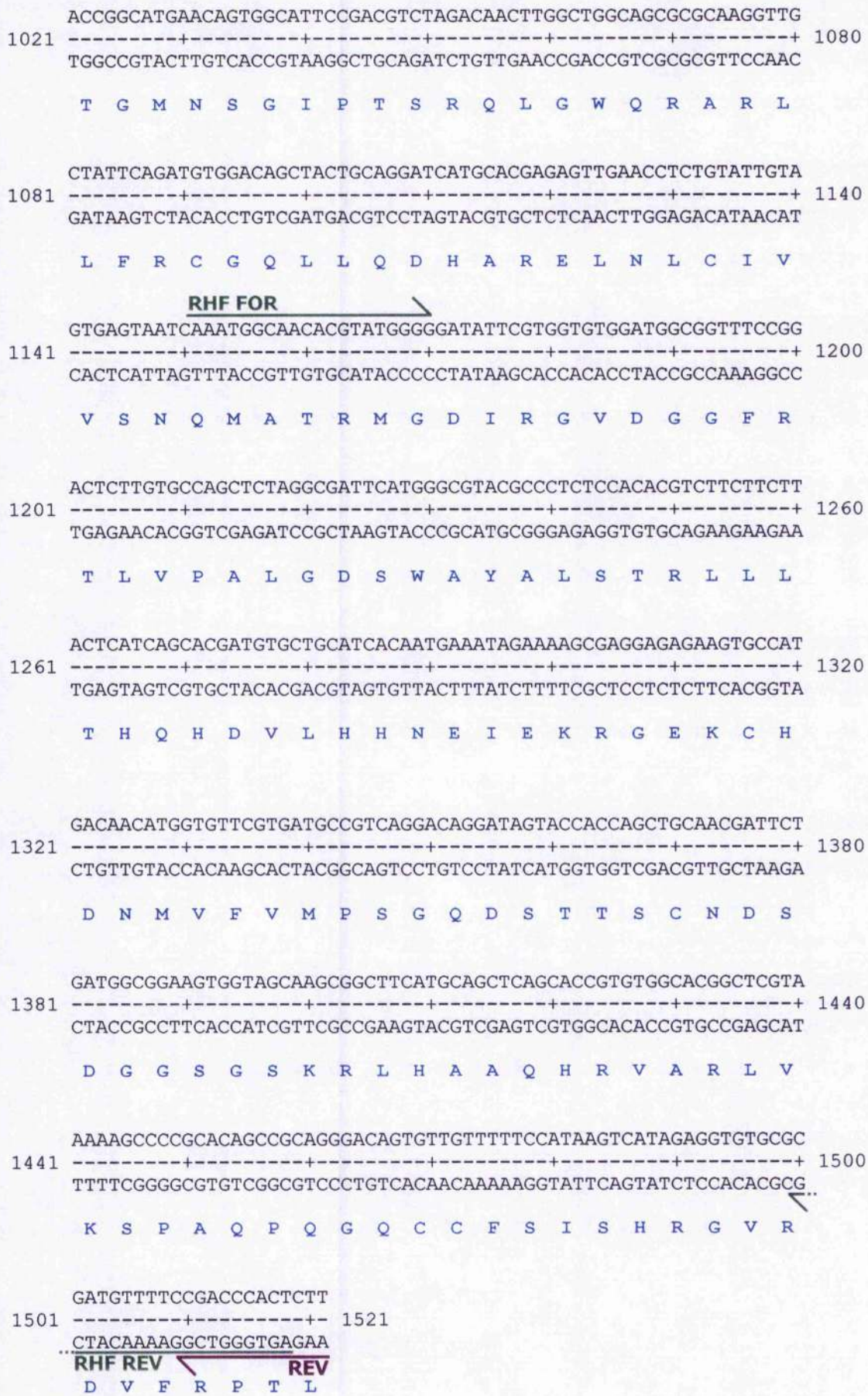
The accession numbers for the Rad51 and RecA genes used during homology and phylogenetic analysis. The sequences were obtained from *Trypanasoma brucei* (T.b), *Leishmania major* (L.m), *Plasmodium falciparum* (P.f), *Trypanasoma cruzi* (T.c), *Escherichia coli* (E.c), *Campylobacter jejuni* (C.j), *Neisseria gonorrhoeae* (N.g), *Bacillus subtilis* (B.s), *Streptomyces lividans* (S.l), *Staphylococcus aureus* (S.a), *Ustilago maydis* (U.m), *Homo sapiens* (H.s), *Saccharomyces cerevisiae* (S.c), *Schizosaccharomyces pombe* (S.p), *Arabidopsis thaliana* (A.t), *Oryza sativa* (O.s) and *Drosophila melanogaster* (D.m). *: Depicts which of the two near identical copies were included in the phylogenetics. ∞: Sequence was unobtainable therefore was not included in the phylogenetics.

Appendix 3: The gene sequence of *RAD51-3*

The RAD51-3 specific primers and the restriction site of the enzyme used during copy number analysis are shown.

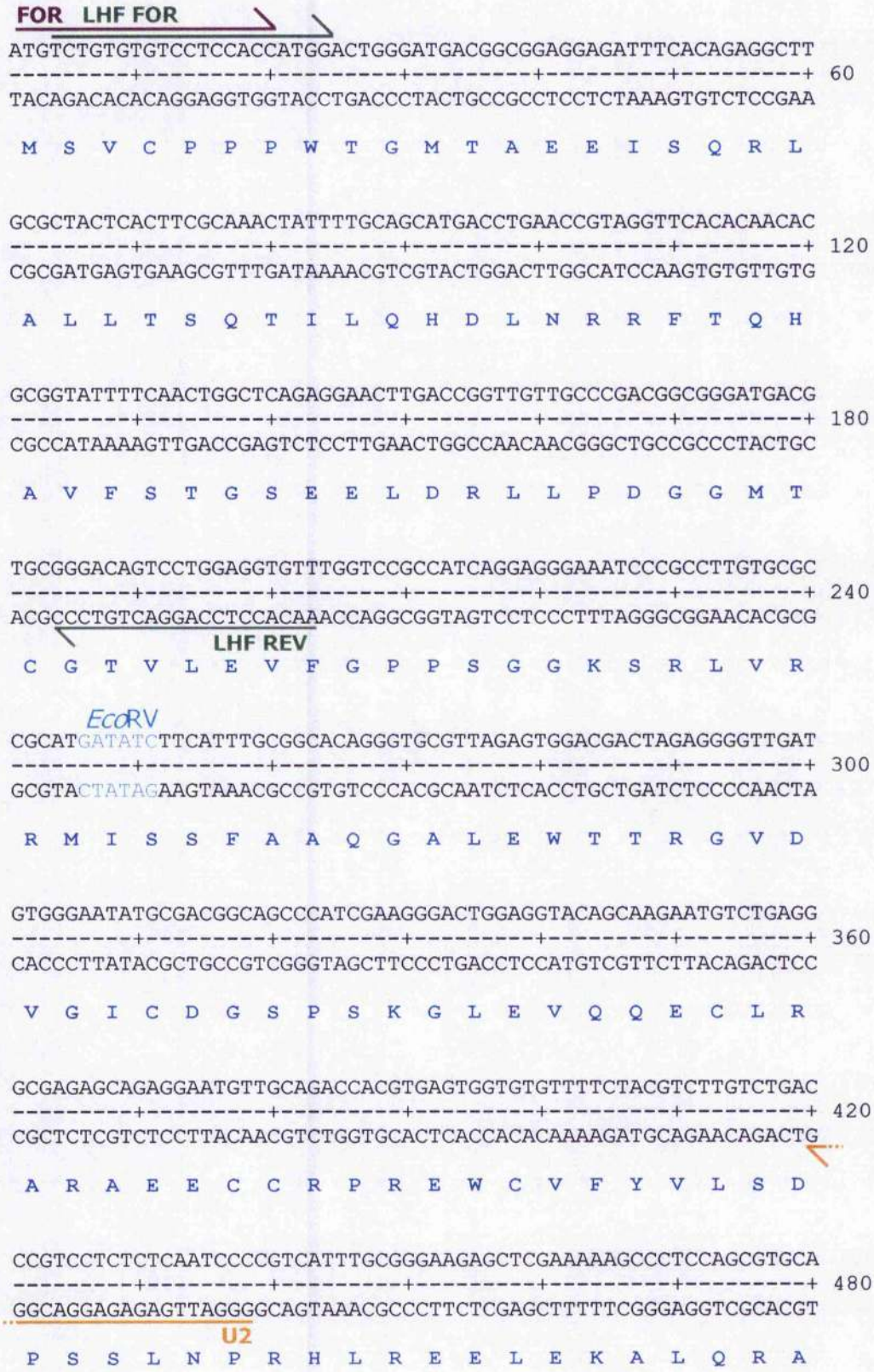



TGTTTGTTTATTGACACGGAGGGGAGCTTTTTACCTGAACGGTTTCGGGAGATTGCATCA
481 540
ACAAACAAATAACTGTGCCTCCCCTCGAAAAATGGACTTGCCAAAGCCCTCTAACGTAGT
C L F I D T E G S F L P E R F R E I A S
GCTGCCGTGGGGCATGTGAGGGAAATCGTACTACAACGTGAGAAAGAAGGACTCGGAGCA
541 600
CGACGGCACCCCGTACACTCCCTTTAGCATGATGTTGCACTCTTTCTTCCTGAGCCTCGT
A A V G H V R E I V L Q R E K E G L G A
GGTAATGTTGGGGTGACAAACGAGGAAAACGGTGTGAGGCAGGCATCTTTCCTCAATGGT
601 660
CCATTACAACCCCACTGTTTGCTCCTTTTGCCACACTCCGTCCGTAGAAAGGAGTTACCA
G N V G V T N E E N G V R Q A S F L N G
GCGATGAATGAGGTGACAGCAGAGCAGACGTGCACCACGTCCTCAGTAGAATCTCGCAAG
661 720
CGCTACTTACTCCACTGTCGTCTCGTCTGCACGTGGTGCAGGAGTCATCTTAGAGCGTTC
A M N E V T A E Q T C T T S S V E S R K
AGGGGAAGGGTTGAGGCTGCGGTCCCACCCCCGCTTGGTGCAATAGTTGGCTCTTTCACA
721 780
TCCCCTTCCCAACTCCGACGCCAGGGTGGGGGCGAACCACGTTATCAACCGAGAAAGTGT
R G R V E A A V P P P L G A I V G S F T
GTAGATTATATTCTTCAGCGAACACAGTACGTTCGCGTGTTGGATGTAGTGTCTCTTATG
781 840
CATCTAATATAAGAAGTCGCTTGTGTCATGCAAGCGCACAACCTACATCACAGAGAATAC
V D Y I L Q R T Q Y V R V L D V V S L M
BamHI
GCACTGCTGAACGGACTTCCCGCGTACATTGCCTCCCACCCAGGGATCCGAATGGTGATC
841 900
CGTGACGACTTGCCTGAAGGGCGCATGTAACGGAGGGTGGGTCCCTAGGCTTACCACTAG
A L L N G L P A Y I A S H P G I R M V I
ATCGACTCTATCGCATTTCCCTTTCGCTCTTTATTGCAGTTGAACGTCAACTCCAACAGT
901 960
TAGCTGAGATAGCGTAAAGGGAAAGCGAGAAATAACGTCAACTTGCAGTTGAGGTTGTCA
I D S I A F P F R S L L Q L N V N S N S
GGTGGCGTTACTGAAAGTCACGGGGCGGTGGGCAAGAGCGTGGCCGGCGTTGATGACTCA
961 1020
CCACCGCAATGACTTTCAGTGCCCCGCCACCCGTTCTCGCACCGGCCGCAACTACTGAGT
KO 3'
G G V T E S H G A V G K S V A G V D D S


ACCGGCATGAACAGTGGCATTCCGACGTCTAGACAACTTGGCTGGCAGCGCGCAAGGTTG
1021 ---------+---------+---------+---------+---------+---------+ 1080
TGGCCGTACTTGTCACCGTAAGGCTGCAGATCTGTTGAACCGACCGTCGCGCGTTCCAAC
T G M N S G I P T S R Q L G W Q R A R L
CTATTCAGATGTGGACAGCTACTGCAGGATCATGCACGAGAGTTGAACCTCTGTATTGTA
1081 ---------+---------+---------+---------+---------+---------+ 1140
GATAAGTCTACACCTGTCGATGACGTCCTAGTACGTGCTCTCAACTTGGAGACATAACAT
L F R C G Q L L Q D H A R E L N L C I V
RHF FOR
GTGAGTAATCAAATGGCAACACGTATGGGGGATATTCGTGGTGTGGATGGCGGTTTCCGG
1141 ---------+---------+---------+---------+---------+---------+ 1200
CACTCATTAGTTTACCGTTGTGCATACCCCCTATAAGCACCACACCTACCGCCAAAGGCC
V S N Q M A T R M G D I R G V D G G F R
ACTCTTGTGCCAGCTCTAGGCGATTCATGGGCGTACGCCCTCTCCACACGTCTTCTTCTT
1201 ---------+---------+---------+---------+---------+---------+ 1260
TGAGAACACGGTCGAGATCCGCTAAGTACCCGCATGCGGGAGAGGTGTGCAGAAGAAGAA
T L V P A L G D S W A Y A L S T R L L L
ACTCATCAGCACGATGTGCTGCATCACAATGAAATAGAAAAGCGAGGAGAGAAGTGCCAT
1261 ---------+---------+---------+---------+---------+---------+ 1320
TGAGTAGTCGTGCTACACGACGTAGTGTTACTTTATCTTTTCGCTCCTCTCTTCACGGTA
T H Q H D V L H H N E I E K R G E K C H
GACAACATGGTGTTCGTGATGCCGTCAGGACAGGATAGTACCACCAGCTGCAACGATTCT
1321 ---------+---------+---------+---------+---------+---------+ 1380
CTGTTGTACCACAAGCACTACGGCAGTCCTGTCCTATCATGGTGGTCGACGTTGCTAAGA
D N M V F V M P S G Q D S T T S C N D S
GATGGCGGAAGTGGTAGCAAGCGGCTTCATGCAGCTCAGCACCGTGTGGCACGGCTCGTA
1381 ---------+---------+---------+---------+---------+---------+ 1440
CTACCGCCTTCACCATCGTTCGCCGAAGTACGTCGAGTCGTGGCACACCGTGCCGAGCAT
D G G S G S K R L H A A Q H R V A R L V
AAAAGCCCCGCACAGCCGCAGGGACAGTGTTGTTTTTCCATAAGTCATAGAGGTGTGCGC
1441 ---------+---------+---------+---------+---------+---------+ 1500
TTTTCGGGGCGTGTCGGCGTCCCTGTCACAACAAAAAGGTATTCAGTATCTCCACACGCG
K S P A Q P Q G Q C C F S I S H R G V R
GATGTTTTCCGACCCACTCTT
1501 ---------+---------+- 1521
CTACAAAAGGCTGGGTGAGAA
RHF REV
REV
D V F R P T L

Appendix 4: The gene sequence of *RAD51-5*

The RAD51-5 specific primers and the restriction site of the enzyme used during copy number analysis are shown.

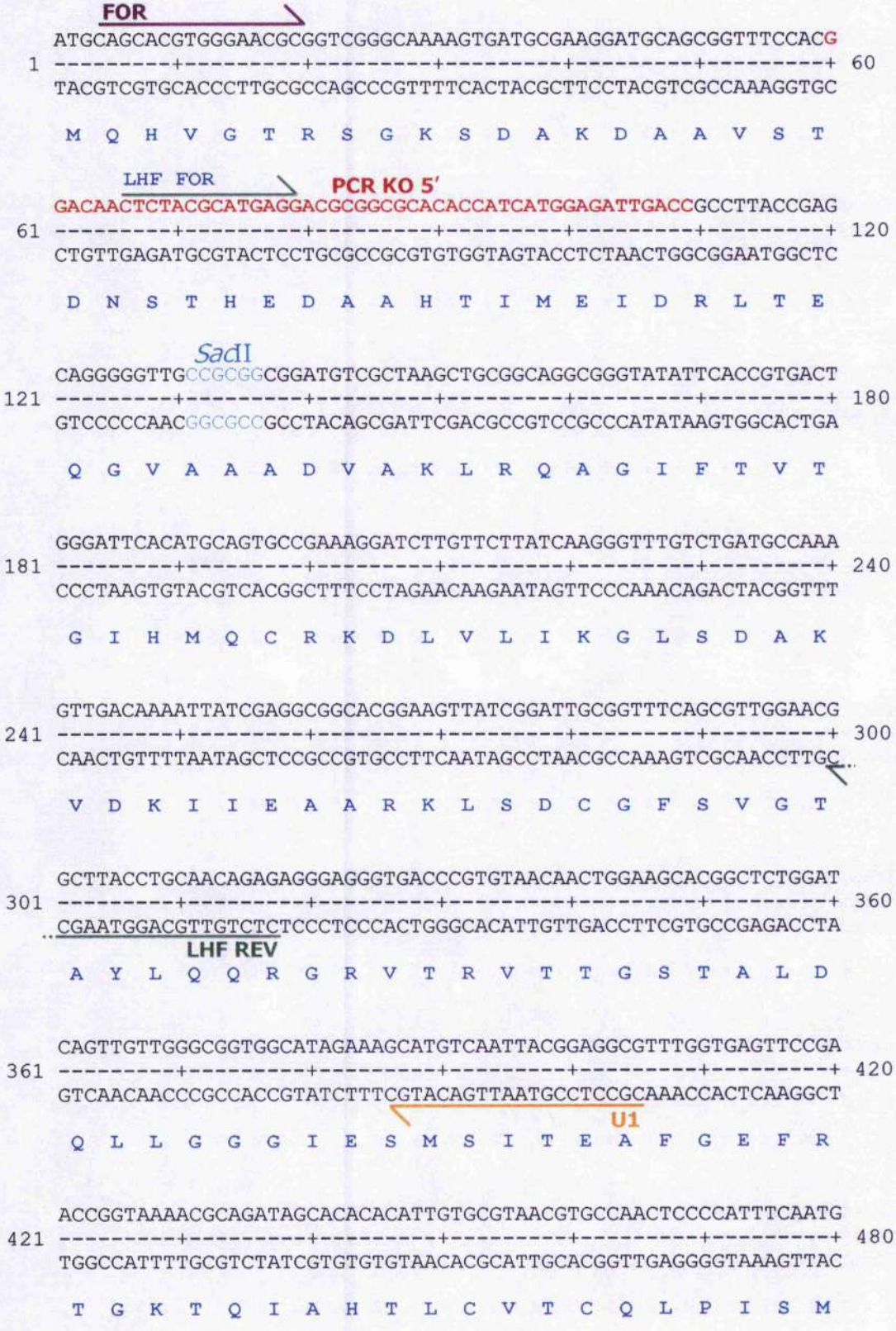



ATGTTGTGTGATGTGCACTGCGAGGGTGAGATTGAAAGGCTCCTCGATGACATCATAGGA
481 540
TACAACACACTACACGTGACGCTCCCACTCTAACTTTCCGAGGAGCTACTGTAGTATCCT
M L C D V H C E G E I E R L L D D I I G
AGGGTACAAGTGGTGAACTTCGCAACCCTAAATGACCTTTTAAACTTCTTCCGTTTTCTG
541 600
TCCCATGTTCACCACTTGAAGCGTTGGGATTTACTGGAAAATTTGAAGAAGGCAAAAGAC
R V Q V V N F A T L N D L L N F F R F L
TACCACGAGGAGTATCCGTCGTCTATGCATCATGCTAACCAACGCCGCGCGTTGGTCGTT
601 660
ATGGTGCTCCTCATAGGCAGCAGATACGTAGTACGATTGGTTGCGGCGCGCAACCAGCAA
Y H E E Y P S S M H H A N Q R R A L V V
ATCGACAGTGTTGCTCGCCTTTGGGATCACCCGACATGTGGAGCAACAAAGCATGCGCGG
661 720
TAGCTGTCACAACGAGCGGAAACCCTAGTGGGCTGTACACCTCGTTGTTTCGTACGCGCC
I D S V A R L W D H P T C G A T K H A R
GACTGGGCGGCGGCTGAACTGGTGCGTGAGTTGCGCAATGTTATTATGCTTGGGAACGGT
721 780
CTGACCCGCCGCCGACTTGACCACGCACTCAACGCGTTACAATAATACGAACCCTTGCCA
D W A A A E L V R E L R N V I M L G N G
TGGCGTGGTGAGTATTGTGGTAACATTGACTCGCATCACTTAGATACCGAAGCTGGGAAC
781 840
ACCGCACCACTCATAACACCATTGTAACTGAGCGTAGTGAATCTATGGCTTCGACCCTTG
W R G E Y C G N I D S H H L D T E A G N
RHF FOR
GTTTTTTGTAGCGTCCAGGCTGCCAATGCGGGTACCTCTGTTGGTGGTAGTGTGGCGGTT
841 900
CAAAAAACATCGCAGGTCCGACGGTTACGCCCATGGAGACAACCACCATCACACCGCCAA
V F C S V Q A A N A G T S V G G S V A V
GTTCTCGTTAACGGGTGCACGAACACGCGTTACTGCCACACCGTTGTTGGACCTTTGGGG
901 960
CAAGAGCAATTGCCCACGTGCTTGTGCGCAATGACGGTGTGGCAACAACCTGGAAACCCC
V L V N G C T N T R Y C H T V V G P L G
GTGCCCTTGTGGTTGGCAGCTGCAGCGGATATTCGTCTTTTCATTGAACCCACATCAACT
961 1020
CACGGGAACACCAACCGTCGACGTCGCCTATAAGCAGAAAAGTAACTTGGGTGTAGTTGA
V P L W L A A A A D I R L F I E P T S T

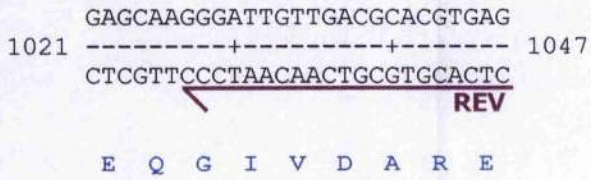

TCGGGTGATTATCCGCCCACTACCGGCGAGGATGACGATGGCATGTTATGCACTCACCAA
1021 ---------+---------+---------+---------+---------+---------+ 1080
AGCCCACTAATAGGCGGGTGATGGCCGCTCCTACTGCTACCGTACAATACGTGAGTGGTT
S G D Y P P T T G E D D D G M L C T H Q
TACGGGGCTGCAAAACACGTATTGGCGGTGCGTGTAGCGAAGGGCAGTGGTTCCTCGGCT
1081 ---------+---------+---------+---------+---------+---------+ 1140
ATGCCCCGACGTTTTGTGCATAACCGCCACGCACATCGCTTCCCGTCACCAAGGAGCCGA
RHF REV
Y G A A K H V L A V R V A K G S G S S A
CCCCGCGTGGGAAACATCTTTTTACCC
1141 ---------+---------+------- 1167
GGGGCGCACCCTTTGTAGAAAAATGGG
REV
P R V G N I F L P

Appendix 5: The gene sequence of *DMC1*

The DMC1 specific primers and the restriction sites of the enzymes used during copy number analysis are shown.




D3
GGGGGGGGCAATGGTAAAGCTATTTATGTTGACACAGAGGCAACTTTCAGACCTGAGCGG
481 ---------+---------+---------+---------+---------+---------+ 540
CCCCCCCCGTTACCATTTCGATAAATACAACTGTGTCTCCGTTGAAAGTCTGGACTCGCC
G G G N G K A I Y V D T E A T F R P E R
ATTAAGCCCATCGCAGAGCGTTTTGGTCTCGACGTTGAGGCCGTACTGGGGAACATTCTC
541 ---------+---------+---------+---------+---------+---------+ 600
TAATTCGGGTAGCGTCTCGCAAAACCAGAGCTGCAACTCCGGCATGACCCCTTGTAAGAG
I K P I A E R F G L D V E A V L G N I L
D2
GTGGCACGAGCTTACACGCACGAGCATCAGATGCACCTCTTGTCAATGGTTGCTGCGAAG
601 ---------+---------+---------+---------+---------+---------+ 660
CACCGTGCTCGAATGTGCGTGCTCGTAGTCTACGTGGAGAACAGTTACCAACGACGCTTC
V A R A Y T H E H Q M H L L S M V A A K
D1
ATGGTAGAGGATCAGTTCAGTCTCCTTGTTGTAGACAGTGTAACGGCTCTCTTTCGCGTG
661 ---------+---------+---------+---------+---------+---------+ 720
TACCATCTCCTAGTCAAGTCAGAGGAACAACATCTGTCACATTGCCGAGAGAAAGCGCAC
M V E D Q F S L L V V D S V T A L F R V
RHF FOR
GATTTTTCTGGCAGAGGGGAGCTCGCGGAGCGGCAGCAGAAACTGGCCAAGATGTTGAGC
721 ---------+---------+---------+---------+---------+---------+ 780
CTAAAAAGACCGTCTCCCCTCGAGCGCCTCGCCGTCGTCTTTGACCGGTTCTACAACTCG
D F S G R G E L A E R Q Q K L A K M L S
AACATGATTAAACTTGCGGAGGAGTATAACGTTGCTGTGTACATAACAAATCAGGTGGTT
781 ---------+---------+---------+---------+---------+---------+ 840
TTGTACTAATTTGAACGCCTCCTCATATTGCAACGACACATGTATTGTTTAGTCCACCAA
N M I K L A E E Y N V A V Y I T N Q V V
BamHI
GCGGATCCTGGTGGCGCCTCCATGTTTGTGGCCGACCCCAAAAAGCCTATTGGAGGTCAT
841 ---------+---------+---------+---------+---------+---------+ 900
CGCCTAGGACCACCGCGGAGGTACAAACACCGGCTGGGGTTTTTCGGATAACCTCCAGTA
A D P G G A S M F V A D P K K P I G G H
ATTCTTGCGCATGCCTCAACGACGCGACTGTCGCTGCGCAAAGGACGGGGGGATCAGCGC
901 ---------+---------+---------+---------+---------+---------+ 960
TAAGAACGCGTACGGAGTTGCTGCGCTGACAGCGACGCGTTTCCTGCCCCCCTAGTCGCG
PCR KO 3'
I L A H A S T T R L S L R K G R G D Q R
GTGTGTAAGATATATGACAGTCCGTCGTTGCCGGAGGTGGAATGTGTTTTCAGCATATCG
961 ---------+---------+---------+---------+---------+---------+ 1020
CACACATTCTATATACTGTCAGGCAGCAACGGCCTCCACCTTACACAAAAGTCGTATAGC
RHF REV
V C K I Y D S P S L P E V E C V F S I S


GAGCAAGGGATTGTTGACGCACGTGAG
1021 ---------+---------+------- 1047
CTCGTTCCCTAACAACTGCGTGCACTC
REV
E Q G I V D A R E

Adzuma, K. (1998). No sliding during homology search by RecA protein. *J.Biol.Chem.* **273**, 31565-31573.

Alarcon, C. M., Son, H. J., Hall, T. and Donelson, J. E. (1994). A monocistronic transcript for a trypanosome variant surface glycoprotein. *Mol.Cell Biol.* **14**, 5579-5591.

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.

Alexiadis, V., Lusser, A. and Kadonaga, J. T. (2004). A conserved N-terminal motif in Rad54 is important for chromatin remodeling and homologous strand pairing. *J.Biol.Chem.* **279**, 27824-27829.

Allred, D. R. and Al Khedery, B. (2004). Antigenic variation and cytoadhesion in *Babesia bovis* and *Plasmodium falciparum*: different logics achieve the same goal. *Mol.Biochem.Parasitol.* **134**, 27-35.

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.

Anderson, D. G. and Kowalczykowski, S. C. (1997). The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a chi-regulated manner. *Cell* **90**, 77-86.

Ansorge, I., Steverding, D., Melville, S., Hartmann, C. and Clayton, C. (1999). Transcription of 'inactive' expression sites in African trypanosomes leads to expression of multiple transferrin receptor RNAs in bloodstream forms. *Mol.Biochem.Parasitol.* **101**, 81-94.

Appelmelk, B. J. and Vandenbrouck-Grauls, C. M. J. E. (2003). Phase Variation in *Helicobacter pylori* Lipopolysaccharide. In 'Antigenic Variation'. (A. Craig and A. Scherf Eds.) pp. 122-141. (Elsevier Ltd.: London.)

Arbel, A., Zenvirth, D. and Simchen, G. (1999). Sister chromatid-based DNA repair is mediated by RAD54, not by DMC1 or TID1. *EMBO J.* **18**, 2648-2658.

Barrett, M. P., Burchmore, R. J., Stich, A., Lazzari, J. O., Frasch, A. C., Cazzulo, J. J. and Krishna, S. (2003). The trypanosomiases. *Lancet* **362**, 1469-1480.

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

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

Barry, J. D. and McCulloch, R. (2003). Trypanopsome Antigenic Variation - a Heavy Investment in the Evasion of Immunity. In 'Antigenic Variation'. (A. Craig and A. Scherf Eds.) pp. 224-242. (Elsevier Ltd.: London.)

Basile, G., Aker, M. and Mortimer, R. K. (1992). Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene RAD51. *Mol.Cell Biol.* **12**, 3235-3246.

Baumann, P. and West, S. C. (1998). DNA end-joining catalyzed by human cell-free extracts. *Proc.Natl.Acad.Sci.U.S.A* **95**, 14066-14070.

Bayles, K. W., Brunskill, E. W., Iandolo, J. J., Hruska, L. L., Huang, S., Pattee, P. A., Smiley, B. K. and Yasbin, R. E. (1994). A genetic and molecular characterization of the recA gene from *Staphylococcus aureus*. *Gene* **147**, 13-20.

Bell, J. S., Harvey, T. I., Sims, A. M. and McCulloch, R. (2004). Characterization of components of the mismatch repair machinery in *Trypanosoma brucei*. *Mol.Microbiol.* **51**, 159-173.

Bell, J. S. and McCulloch, R. (2003). Mismatch repair regulates homologous recombination, but has little influence on antigenic variation, in *Trypanosoma brucei*. *J.Biol.Chem.* **278**, 45182-45188.

Bennett, R. L. and Holloman, W. K. (2001). A RecA homologue in *Ustilago maydis* that is distinct and evolutionarily distant from Rad51 actively promotes DNA pairing reactions in the absence of auxiliary factors. *Biochemistry* **40**, 2942-2953.

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.

Berriman, M., Hall, N., Sheader, K., Bringaud, F., Tiwari, B., Isobe, T., Bowman, S., Corton, C., Clark, L., Cross, G. A., Hoek, M., Zanders, T., Berberof, M., Borst, P. and Rudenko, G. (2002). The architecture of variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **122**, 131-140.

Bhattacharyya, M. K. and Kumar, N. (2003). Identification and molecular characterisation of DNA damaging agent induced expression of *Plasmodium falciparum* recombination protein PfRad51. *Int.J.Parasitol.* **33**, 1385-1392.

Bingle, L. E., Eastlake, J. L., Bailey, M. and Gibson, W. C. (2001). A novel GFP approach for the analysis of genetic exchange in trypanosomes allowing the in situ detection of mating events. *Microbiology* **147**, 3231-3240.

Bishop, D. K. (1994). RecA homologs Dmc1 and Rad51 interact to form multiple nuclear complexes prior to meiotic chromosome synapsis. *Cell* **79**, 1081-1092.

Bishop, D. K., Ear, U., Bhattacharyya, A., Calderone, C., Beckett, M., Weichselbaum, R. R. and Shinohara, A. (1998). Xrcc3 is required for assembly of Rad51 complexes in vivo. *J.Biol.Chem.* **273**, 21482-21488.

Bishop, D. K., Park, D., Xu, L. and Kleckner, N. (1992). DMC1: a meiosis-specific yeast homolog of E. coli recA required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* **69**, 439-456.

Black, C. G., Fyfe, J. A. and Davies, J. K. (1998). Absence of an SOS-like system in *Neisseria gonorrhoeae*. *Gene* **208**, 61-66.

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.

Boddy, M. N., Gaillard, P. H., McDonald, W. H., Shanahan, P., Yates, J. R., III and Russell, P. (2001). Mus81-Eme1 are essential components of a Holliday junction resolvase. *Cell* **107**, 537-548.

Boothroyd, J. C., Cross, G. A., Hoeijmakers, J. H. and Borst, P. (1980). A variant surface glycoprotein of *Trypanosoma brucei* synthesized with a C-terminal hydrophobic 'tail' absent from purified glycoprotein. *Nature* **288**, 624-626.

Borst, P. (2002). Antigenic Variation in Eukaryotic Parasites. In 'Mobile DNA II'. (N. L. Craig and D. E. Berg Eds.) pp. 953-971. (ASM Press: Washington.)

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

Borst, P., Rudenko, G., Taylor, M. C., Blundell, P. A., van Leeuwen, F., Bitter, W., Cross, M., and McCulloch, R. (1996). Antigenic variation in trypanosomes. *Arch.Med.Res.* **27**, 379-388.

Brendel, V., Brocchieri, L., Sandler, S. J., Clark, A. J. and Karlin, S. (1997). Evolutionary comparisons of RecA-like proteins across all major kingdoms of living organisms. *J.Mol.Evol.* **44**, 528-541.

Burton, P. (2003) PhD Thesis. The University of Glasgow.

Campbell, C. and Romero, D. P. (1998). Identification and characterization of the RAD51 gene from the ciliate *Tetrahymena thermophila*. *Nucleic Acids Res.* **26**, 3165-3172.

Campbell, D. A., Thomas, S. and Sturm, N. R. (2003). Transcription in kinetoplastid protozoa: why be normal? *Microbes.Infect.* **5**, 1231-1240.

Cao, L., Alani, E. and Kleckner, N. (1990). A pathway for generation and processing of double-strand breaks during meiotic recombination in *S. cerevisiae*. *Cell* **61**, 1089-1101.

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

Cary, R. B., Peterson, S. R., Wang, J., 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.U.S.A* **94**, 4267-4272.

Chattopadhyay, A., Jones, N. G., Nietlispach, D., Nielsen, P. R., Voorheis, H. P., Mott, H. R. and Carrington, M. (2005). Structure of the C-terminal Domain from *Trypanosoma brucei* Variant Surface Glycoprotein MITat1.2. *J.Biol.Chem.* **280**, 7228-7235.

Chaves, I., Rudenko, G., Dirks-Mulder, A., Cross, M. and Borst, P. (1999). Control of variant surface glycoprotein gene-expression sites in *Trypanosoma brucei*. *EMBO J.* **18**, 4846-4855.

Chaves, I., Zomerdijk, J., Dirks-Mulder, A., Dirks, R. W., Raap, A. K. and Borst, P. (1998). Subnuclear localization of the active variant surface glycoprotein gene expression site in *Trypanosoma brucei*. *Proc.Natl.Acad.Sci.U.S.A* **95**, 12328-12333.

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

Chen, P. L., Chen, C. F., Chen, Y., Xiao, J., Sharp, Z. D. and Lee, W. H. (1998). The BRC repeats in BRCA2 are critical for RAD51 binding and resistance to methyl methanesulfonate treatment. *Proc.Natl.Acad.Sci.U.S.A* **95**, 5287-5292.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **31**, 3497-3500.

Connelly, J. C., de Leau, E. S. and Leach, D. R. (1999). DNA cleavage and degradation by the SbcCD protein complex from *Escherichia coli*. *Nucleic Acids Res.* **27**, 1039-1046.

Connolly, B., Parsons, C. A., Benson, F. E., Dunderdale, H. J., Sharples, G. J., Lloyd, R. G. and West, S. C. (1991). Resolution of Holliday junctions in vitro requires the *Escherichia coli* ruvC gene product. *Proc.Natl.Acad.Sci.U.S.A* **88**, 6063-6067.

Constantinou, A., Chen, X. B., McGowan, C. H. and West, S. C. (2002). Holliday junction resolution in human cells: two junction endonucleases with distinct substrate specificities. *EMBO J.* **21**, 5577-5585.

Conway, A. B., Lynch, T. W., Zhang, Y., Fortin, G. S., Fung, C. W., Symington, L. S. and Rice, P. A. (2004). Crystal structure of a Rad51 filament. *Nat.Struct.Mol.Biol.* **11**, 791-796.

Conway, C., McCulloch, R., Ginger, M. L., Robinson, N. P., Browitt, A. and Barry, J. D. (2002a). Ku is important for telomere maintenance, but not for differential expression of telomeric VSG genes, in African trypanosomes. *J.Biol.Chem.* **277**, 21269-21277.

Conway, C., Proudfoot, C., Burton, P., Barry, J. D. and McCulloch, R. (2002b). Two pathways of homologous recombination in *Trypanosoma brucei*. *Mol.Microbiol.* **45**, 1687-1700.

Cox, M. M., Goodman, M. F., Kreuzer, K. N., Sherratt, D. J., Sandler, S. J. and Marians, K. J. (2000). The importance of repairing stalled replication forks. *Nature* **404**, 37-41.

Craig, A. and Scherf, A. (2003). Antigenic Variation. (Elselvier Ltd.: London.)

Cromie, G. A., Connelly, J. C. and Leach, D. R. (2001). Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. *Mol.Cell* **8**, 1163-1174.

Cross, M., Kieft, R., Sabatini, R., Dirks-Mulder, A., Chaves, I. and Borst, P. (2002). J-binding protein increases the level and retention of the unusual base J in trypanosome DNA. *Mol.Microbiol.* **46**, 37-47.

Cross, M., Kieft, R., Sabatini, R., Wilm, M., de Kort, M., van der Marel, G. A., van Boom, J. H., van Leeuwen, F. and Borst, P. (1999). The modified base J is the target for a novel DNA-binding protein in kinetoplastid protozoans. *EMBO J.* **18**, 6573-6581.

D'Amours, D. and Jackson, S. P. (2002). The Mre11 complex: at the crossroads of dna repair and checkpoint signalling. *Nat.Rev.Mol.Cell Biol.* **3**, 317-327.

D'Orso, I., De Gaudenzi, J. G. and Frasch, A. C. (2003). RNA-binding proteins and mRNA turnover in trypanosomes. *Trends Parasitol.* **19**, 151-155.

Dalgaard, J. Z. and Klar, A. J. (2001). Does *S. pombe* exploit the intrinsic asymmetry of DNA synthesis to imprint daughter cells for mating-type switching? *Trends Genet.* **17**, 153-157.

Datta, A., Adjiri, A., New, L., Crouse, G. F. and Jinks, Robertson S. (1996). Mitotic crossovers between diverged sequences are regulated by mismatch repair proteins in *Saccaromyces cerevisiae*. *Mol.Cell Biol.* **16**, 1085-1093.

Davies, A. A., Masson, J. Y., McIlwraith, M. J., Stasiak, A. Z., Stasiak, A., Venkitaraman, A. R. and West, S. C. (2001). Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol.Cell* **7**, 273-282.

Davis, A. P. and Symington, L. S. (2004). RAD51-dependent break-induced replication in yeast. *Mol.Cell Biol.* **24**, 2344-2351.

De Greef, C. and Hamers, R. (1994). The serum resistance-associated (SRA) gene of *Trypanosoma brucei rhodesiense* encodes a variant surface glycoprotein-like protein. *Mol.Biochem.Parasitol.* **68**, 277-284.

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

Deitsch, K. W. and Hviid, L. (2004). Variant surface antigens, virulence genes and the pathogenesis of malaria. *Trends Parasitol.* **20**, 562-566.

Ding, Z., Wang, T., Chong, K. and Bai, S. (2001). Isolation and characterization of OsDMC1, the rice homologue of the yeast DMC1 gene essential for meiosis. *Sex.Plant Reprod.* **13**, 285-288.

DiRuggiero, J., Brown, J. R., Bogert, A. P. and Robb, F. T. (1999). DNA repair systems in archaea: mementos from the last universal common ancestor? *J.Mol.Evol.* **49**, 474-484.

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

Donelson, J. E., Murphy, W. J., Brentano, S. T., Rice-Ficht, A. C. and Cain, G. D. (1983). Comparison of the expression-linked extra copy (ELC) and basic copy (BC) genes of a trypanosome surface antigen. *J.Cell Biochem.* **23**, 1-12.

Dooijes, D., Chaves, I., Kieft, R., Dirks-Mulder, A., Martin, W. and Borst, P. (2000). Base J originally found in kinetoplastida is also a minor constituent of nuclear DNA of Euglena gracilis. *Nucleic Acids Res.* **28**, 3017-3021.

Doutriaux, M. P., Couteau, F., Bergounioux, C. and White, C. (1998). Isolation and characterisation of the RAD51 and DMC1 homologs from *Arabidopsis thaliana*. *Mol.Gen.Genet.* **257**, 283-291.

Dujon, B. (1989). Group I introns as mobile genetic elements: facts and mechanistic speculations--a review. *Gene* **82**, 91-114.

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

Elliott, B. and Jasin, M. (2001). Repair of double-strand breaks by homologous recombination in mismatch repair-defective mammalian cells. *Mol.Cell Biol.* **21**, 2671-2682.

Erck, C., Frank, R. and Wehland, J. (2000). Tubulin-tyrosine ligase, a long-lasting enigma. *Neurochem.Res.* **25**, 5-10.

Ersfeld, K., Melville, S. E. and Gull, K. (1999). Nuclear and genome organization of *Trypanosoma brucei*. *Parasitol.Today* **15**, 58-63.

Essers, J., Houtsmuller, A. B., van Veelen, L., Paulusma, C., Nigg, A. L., Pastink, A., Vermeulen, W., Hoeijmakers, J. H. and Kanaar, R. (2002). Nuclear dynamics of RAD52 group homologous recombination proteins in response to DNA damage. *EMBO J.* **21**, 2030-2037.

Evans, E. and Alani, E. (2000). Roles for mismatch repair factors in regulating genetic recombination. *Mol.Cell Biol.* **20**, 7839-7844.

Ferguson, D. O. and Holloman, W. K. (1996). Recombinational repair of gaps in DNA is asymmetric in *Ustilago maydis* and can be explained by a migrating D-loop model. *Proc.Natl.Acad.Sci.U.S.A* **93**, 5419-5424.

Ferguson, D. O., Rice, M. C., Rendi, M. H., Kotani, H., Kmiec, E. B. and Holloman, W. K. (1997). Interaction between *Ustilago maydis* REC2 and RAD51 genes in DNA repair and mitotic recombination. *Genetics* **145**, 243-251.

Forget, A. L., Bennett, B. T. and Knight, K. L. (2004). Xrcc3 is recruited to DNA double strand breaks early and independent of Rad51. *J.Cell Biochem.* **93**, 429-436.

Forsburg, S. L. (2004). Eukaryotic MCM proteins: beyond replication initiation. *Microbiol.Mol.Biol.Rev.* **68**, 109-31, table.

Fortin, G. S. and Symington, L. S. (2002). Mutations in yeast Rad51 that partially bypass the requirement for Rad55 and Rad57 in DNA repair by increasing the stability of Rad51-DNA complexes. *EMBO J.* **21**, 3160-3170.

Freitas-Junior, L. H., Bottius, E., Pirrit, L. A., Deitsch, K. W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T. E. and Scherf, A. (2000). Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature* **407**, 1018-1022.

Fukushima, K., Tanaka, Y., Nabeshima, K., Yoneki, T., Tougan, T., Tanaka, S. and Nojima, H. (2000). Dmc1 of *Schizosaccharomyces pombe* plays a role in meiotic recombination. *Nucleic Acids Res.* **28**, 2709-2716.

Galinski, M. R. and Corredor, V. (2004). Variant antigen expression in malaria infections: posttranscriptional gene silencing, virulence and severe pathology. *Mol.Biochem.Parasitol.* **134**, 17-25.

Gasior, S. L., Wong, A. K., Kora, Y., Shinohara, A. and Bishop, D. K. (1998). Rad52 associates with RPA and functions with rad55 and rad57 to assemble meiotic recombination complexes. *Genes Dev.* **12**, 2208-2221.

Gatton, M. L., Peters, J. M., Fowler, E. V. and Cheng, Q. (2003). Switching rates of *Plasmodium falciparum var* genes: faster than we thought? *Trends Parasitol.* **19**, 202-208.

Gibson, W. (2001). Sex and evolution in trypanosomes. *Int.J.Parasitol.* **31**, 643-647.

Gibson, W. and Stevens, J. (1999). Genetic exchange in the trypanosomatidae. *Adv.Parasitol.* **43**, 1-46.

Gibson, W. and Whittington, H. (1993). Genetic exchange in *Trypanosoma brucei*: selection of hybrid trypanosomes by introduction of genes conferring drug resistance. *Mol.Biochem.Parasitol.* **60**, 19-26.

Giloni, L., Takeshita, M., Johnson, F., Iden, C. and Grollman, A. P. (1981). Bleomycin-induced strand-scission of DNA. Mechanism of deoxyribose cleavage. *J.Biol.Chem.* **256**, 8608-8615.

Ginger, M. L., Blundell, P. A., Lewis, A. M., Browitt, A., Gunzl, A. and Barry, J. D. (2002). Ex Vivo and In Vitro Identification of a Consensus Promoter for VSG Genes Expressed by Metacyclic-Stage Trypanosomes in the Tsetse Fly. *Eukaryot.Cell* **1**, 1000-1009.

Glaab, W. E., Risinger, J. I., Umar, A., Barrett, J. C., Kunkel, T. A. and Tindall, K. R. (1998). Cellular resistance and hypermutability in mismatch repair-deficient human cancer cell lines following treatment with methyl methanesulfonate. *Mutat.Res.* **398**, 197-207.

Gommers-Ampt, J., Lutgerink, J. and Borst, P. (1991). A novel DNA nucleotide in *Trypanosoma brucei* only present in the mammalian phase of the life-cycle. *Nucleic Acids Res.* **19**, 1745-1751.

Gommers-Ampt, J. H., Teixeira, A. J., Van de, Werken G., van Dijk, W. J. and Borst, P. (1993). The identification of hydroxymethyluracil in DNA of *Trypanosoma brucei*. *Nucleic Acids Res.* **21**, 2039-2043.

Gottesdiener, K. M., Goriparthi, L., Masucci, J. P. and Van der Ploeg, L. H. (1992). A proposed mechanism for promoter-associated DNA rearrangement events at a variant surface glycoprotein gene expression site. *Mol.Cell Biol.* **12**, 4784-4795.

Graham, S. V., Terry, S. and Barry, J. D. (1999). A structural and transcription pattern for variant surface glycoprotein gene expression sites used in metacyclic stage *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **103**, 141-154.

Grishchuk, A. L. and Kohli, J. (2003). Five RecA-like proteins of *Schizosaccharomyces pombe* are involved in meiotic recombination. *Genetics* **165**, 1031-1043.

Guerry, P., Pope, P. M., Burr, D. H., Leifer, J., Joseph, S. W. and Bourgeois, A. L. (1994). Development and characterization of recA mutants of *Campylobacter jejuni* for inclusion in attenuated vaccines. *Infect.Immun.* **62**, 426-432.

Gunzl, A., Bruderer, T., Laufer, G., Schimanski, B., Tu, L. C., Chung, H. M., Lee, P. T. and Lee, M. G. (2003). RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Eukaryot.Cell* **2**, 542-551.

Gupta, R. C., Golub, E., Bi, B. and Radding, C. M. (2001). The synaptic activity of HsDmc1, a human recombination protein specific to meiosis. *Proc.Natl.Acad.Sci.U.S.A* **98**, 8433-8439.

Haaf, T., Golub, E. I., Reddy, G., Radding, C. M. and Ward, D. C. (1995). Nuclear foci of mammalian Rad51 recombination protein in somatic cells after DNA damage and its localization in synaptonemal complexes. *Proc.Natl.Acad.Sci.U.S.A* **92**, 2298-2302.

Haaf, T., Raderschall, E., Reddy, G., Ward, D. C., Radding, C. M. and Golub, E. I. (1999). Sequestration of mammalian Rad51-recombination protein into micronuclei. *J.Cell Biol.* **144**, 11-20.

Haber, J. E. (2002). Switching of Saccharomyces Mating-Type Genes. In 'Mobile DNA II'. (N. L. Craig and D. E. Berg Eds.) pp. 927-953. (ASM Press: Washington.)

Hamatake, R. K., Dykstra, C. C. and Sugino, A. (1989). Presynapsis and synapsis of DNA promoted by the STP alpha and single-stranded DNA-binding proteins from *Saccharomyces cerevisiae*. *J.Biol.Chem.* **264**, 13336-13342.

Hammarsten, O. and Chu, G. (1998). DNA-dependent protein kinase: DNA binding and activation in the absence of Ku. *Proc.Natl.Acad.Sci.U.S.A* **95**, 525-530.

Hartung, F. and Puchta, H. (2000). Molecular characterisation of two paralogous SPO11 homologues in *Arabidopsis thaliana*. *Nucleic Acids Res.* **28**, 1548-1554.

Hays, S. L., Firmenich, A. A. and Berg, P. (1995). Complex formation in yeast double-strand break repair: participation of Rad51, Rad52, Rad55, and Rad57 proteins. *Proc.Natl.Acad.Sci.U.S.A* **92**, 6925-6929.

Hendriks, E., van Deursen, F. J., Wilson, J., Sarkar, M., Timms, M. and Matthews, K. R. (2000). Life-cycle differentiation in *Trypanosoma brucei*: molecules and mutants. *Biochem.Soc.Trans.* **28**, 531-536.

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

Holliday, R. (1964). A mechanism for gene conversion in fungi. *Genet.Res.* **5**, 282-304.

Hollingsworth, N. M., Ponte, L. and Halsey, C. (1995). MSH5, a novel MutS homolog, facilitates meiotic reciprocal recombination between homologs in *Saccharomyces cerevisiae* but not mismatch repair. *Genes Dev.* **9**, 1728-1739.

Holmes, A. M. and Haber, J. E. (1999). Double-strand break repair in yeast requires both leading and lagging strand DNA polymerases. *Cell* **96**, 415-424.

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

Horn, D. and Cross, G. A. (1997). Analysis of *Trypanosoma brucei* vsg expression site switching in vitro. *Mol.Biochem.Parasitol.* **84**, 189-201.

Horrocks, P., Pinches, R., Christodoulou, Z., Kyes, S. A. and Newbold, C. I. (2004). Variable var transition rates underlie antigenic variation in malaria. *Proc.Natl.Acad.Sci.U.S.A* **101**, 11129-11134.

Hunter, N. and Borts, R. H. (1997). Mlh1 is unique among mismatch repair proteins in its ability to promote crossing-over during meiosis. *Genes Dev.* **11**, 1573-1582.

Irmer, H. and Clayton, C. (2001). Degradation of the unstable EP1 mRNA in *Trypanosoma brucei* involves initial destruction of the 3'-untranslated region. *Nucleic Acids Res.* **29**, 4707-4715.

Ivanov, E. L., Sugawara, N., White, C. I., Fabre, F. and Haber, J. E. (1994). Mutations in XRS2 and RAD50 delay but do not prevent mating-type switching in *Saccharomyces cerevisiae*. *Mol.Cell Biol.* **14**, 3414-3425.

Jackson, S. P. (2001). Detecting, signalling and repairing DNA double-strand breaks. *Biochem.Soc.Trans.* **29**, 655-661.

Jang, Y. K., Jin, Y. H., Kim, E. M., Fabre, F., Hong, S. H. and Park, S. D. (1994). Cloning and sequence analysis of rhp51+, a *Schizosaccharomyces pombe* homolog of the *Saccharomyces cerevisiae* RAD51 gene. *Gene* **142**, 207-211.

Janzen, C. J., Lander, F., Dreesen, O. and Cross, G. A. (2004). Telomere length regulation and transcriptional silencing in KU80-deficient *Trypanosoma brucei*. *Nucleic Acids Res.* **32**, 6575-6584.

Jenni, L., Marti, S., Schweizer, J., Betschart, B., Le Page, R. W., Wells, J. M., Tait, A., Paindavoine, P., Pays, E. and Steinert, M. (1986). Hybrid formation between African trypanosomes during cyclical transmission. *Nature* **322**, 173-175.

Jiang, H., Xie, Y., Houston, P., Stemke-Hale, K., Mortensen, U. H., Rothstein, R. and Kodadek, T. (1996). Direct association between the yeast Rad51 and Rad54 recombination proteins. *J.Biol.Chem.* **271**, 33181-33186.

Johnson, R. D. and Symington, L. S. (1995). Functional differences and interactions among the putative RecA homologs Rad51, Rad55, and Rad57. *Mol.Cell Biol.* **15**, 4843-4850.

Johzuka, K. and Ogawa, H. (1995). Interaction of Mre11 and Rad50: two proteins required for DNA repair and meiosis-specific double-strand break formation in Saccharomyces cerevisiae. *Genetics* **139**, 1521-1532.

Kafer, E. (1987). MMS sensitivity of all amino acid-requiring mutants in aspergillus and its suppression by mutations in a single gene. *Genetics* **115**, 671-676.

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**, 33-44.

Kathiresan, A., Khush, G. S. and Bennett, J. (2002). Two rice DMC1 genes are differentially expressed during meiosis and during haploid and diploid mitosis. *Sex.Plant Reprod.* **14**, 257-267.

Keeney, S. (2001). Mechanism and control of meiotic recombination initiation. *Curr.Top.Dev.Biol.* **52**, 1-53.

Keeney, S., Baudat, F., Angeles, M., Zhou, Z. H., Copeland, N. G., Jenkins, N. A., Manova, K. and Jasin, M. (1999). A mouse homolog of the *Saccharomyces cerevisiae* meiotic recombination DNA transesterase Spo11p. *Genomics* **61**, 170-182.

Keeney, S., Giroux, C. N. and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88**, 375-384.

Kleckner, N. (1996). Meiosis: how could it work? *Proc.Natl.Acad.Sci.U.S.A* **93**, 8167-8174.

Klein, H. L. (1997). RDH54, a RAD54 homologue in *Saccharomyces cerevisiae*, is required for mitotic diploid-specific recombination and repair and for meiosis. *Genetics* **147**, 1533-1543.

Kojic, M., Kostrub, C. F., Buchman, A. R. and Holloman, W. K. (2002). BRCA2 homolog required for proficiency in DNA repair, recombination, and genome stability in *Ustilago maydis*. *Mol.Cell* **10**, 683-691.

Komori, K., Miyata, T., DiRuggiero, J., Holley-Shanks, R., Hayashi, I., Cann, I. K., Mayanagi, K., Shinagawa, H. and Ishino, Y. (2000). Both RadA and RadB are involved in homologous recombination in *Pyrococcus furiosus*. *J.Biol.Chem.* **275**, 33782-33790.

Koomey, J. M. and Falkow, S. (1987). Cloning of the recA gene of *Neisseria gonorrhoeae* and construction of gonococcal recA mutants. *J.Bacteriol.* **169**, 790-795.

Kooter, J. M., Winter, A. J., de Oliveira, C., Wagter, R. and Borst, P. (1988). Boundaries of telomere conversion in *Trypanosoma brucei*. *Gene* **69**, 1-11.

Kowalczykowski, S. C. (2000). Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem.Sci.* **25**, 156-165.

Kowalczykowski, S. C. and Eggleston, A. K. (1994). Homologous pairing and DNA strand-exchange proteins. *Annu.Rev.Biochem.* **63**, 991-1043.

Kraus, E., Leung, W. Y. and Haber, J. E. (2001). Break-induced replication: a review and an example in budding yeast. *Proc.Natl.Acad.Sci.U.S.A* **98**, 8255-8262.

Krogh, B. O. and Symington, L. S. (2004). Recombination Proteins in Yeast. *Annu.Rev.Genet.*

Kuo, M. L., Chou, Y. W., Chau, Y. P. and Huang, T. S. (1997). Resistance to apoptosis induced by alkylating agents in v-Ha-ras-transformed cells due to defect in p53 function. *Mol.Carcinog.* **18**, 221-231.

Kupiec, M. (2000). Damage-induced recombination in the yeast *Saccharomyces cerevisiae*. *Mutat.Res.* **451**, 91-105.

Kuzminov, A. (1999). Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol.Mol.Biol.Rev.* **63**, 751-813, table.

LaCount, D. J. and Donelson, J. E. (2001). RNA interference in African trypanosomes. *Protist.* **152**, 103-111.

Lamont, G. S., Tucker, R. S. and Cross, G. A. (1986). Analysis of antigen switching rates in *Trypanosoma brucei*. *Parasitology* **92 (Pt 2)**, 355-367.

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

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

Lichten, M. (2001). Meiotic recombination: breaking the genome to save it. *Curr.Biol.* **11**, R253-R256.

Lieber, M. R., Ma, Y., Pannicke, U. and Schwarz, K. (2003). Mechanism and regulation of human non-homologous DNA end-joining. *Nat.Rev.Mol.Cell Biol.* **4**, 712-720.

Ligtenberg, M. J., Bitter, W., Kieft, R., Steverding, D., Janssen, H., Calafat, J. and Borst, P. (1994). Reconstitution of a surface transferrin binding complex in insect form *Trypanosoma brucei*. *EMBO J.* **13**, 2565-2573.

Lillico, S., Field, M. C., Blundell, P., Coombs, G. H. and Mottram, J. C. (2003). Essential roles for GPI-anchored proteins in African trypanosomes revealed using mutants deficient in GPI8. *Mol.Biol.Cell* **14**, 1182-1194.

Lio, Y. C., Mazin, A. V., Kowalczykowski, S. C. and Chen, D. J. (2003). Complex formation by the human Rad51B and Rad51C DNA repair proteins and their activities in vitro. *J.Biol.Chem.* **278**, 2469-2478.

Lisby, M., Antunez, de Mayolo, Mortensen, U. H. and Rothstein, R. (2003a). Cell cycle-regulated centers of DNA double-strand break repair. *Cell Cycle* **2**, 479-483.

Lisby, M., Barlow, J. H., Burgess, R. C. and Rothstein, R. (2004). Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* **118**, 699-713.

Lisby, M., Mortensen, U. H. and Rothstein, R. (2003b). Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. *Nat.Cell Biol.*

Lisby, M. and Rothstein, R. (2004). DNA damage checkpoint and repair centers. *Curr.Opin.Cell Biol.* **16**, 328-334.

Lisby, M., Rothstein, R. and Mortensen, U. H. (2001). Rad52 forms DNA repair and recombination centers during S phase. *Proc.Natl.Acad.Sci.U.S.A* **98**, 8276-8282.

Little, J. W. and Mount, D. W. (1982). The SOS regulatory system of *Escherichia coli*. *Cell* **29**, 11-22.

Liu, A. Y., Michels, P. A., Bernards, A. and Borst, P. (1985). Trypanosome variant surface glycoprotein genes expressed early in infection. *J.Mol.Biol.* **182**, 383-396.

Liu, A. Y., Van der Ploeg, L. H., Rijsewijk, F. A. 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.

Liu, N., Schild, D., Thelen, M. P. and Thompson, L. H. (2002). Involvement of Rad51C in two distinct protein complexes of Rad51 paralogs in human cells. *Nucleic Acids Res.* **30**, 1009-1015.

Liu, Y. and Maizels, N. (2000). Coordinated response of mammalian Rad51 and Rad52 to DNA damage. *EMBO Rep.* **1**, 85-90.

Liu, Y., Masson, J. Y., Shah, R., O'Regan, P. and West, S. C. (2004). RAD51C is required for Holliday junction processing in mammalian cells. *Science* **303**, 243-246.

Lo, T., Pellegrini, L., Venkitaraman, A. R. and Blundell, T. L. (2003). Sequence fingerprints in BRCA2 and RAD51: implications for DNA repair and cancer. *DNA Repair (Amst)* **2**, 1015-1028.

Lukas, C., Falck, J., Bartkova, J., Bartek, J. and Lukas, J. (2003). Distinct spatiotemporal dynamics of mammalian checkpoint regulators induced by DNA damage. *Nat.Cell Biol.* **5**, 255-260.

MacLeod, A., Turner, C. M. and Tait, A. (1999). A high level of mixed *Trypanosoma brucei* infections in tsetse flies detected by three hypervariable minisatellites. *Mol.Biochem.Parasitol.* **102**, 237-248.

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**, 7131-7136.

Malkova, A., Naylor, M. L., Yamaguchi, M., Ira, G. and Haber, J. E. (2005). RAD51-dependent break-induced replication differs in kinetics and checkpoint responses from RAD51-mediated gene conversion. *Mol.Cell Biol.* **25**, 933-944.

Malkova, A., Signon, L., Schaefer, C. B., Naylor, M. L., Theis, J. F., Newlon, C. S. and Haber, J. E. (2001). RAD51-independent break-induced replication to repair a broken chromosome depends on a distant enhancer site. *Genes Dev.* **15**, 1055-1060.

Maser, R. S., Monsen, K. J., Nelms, B. E. and Petrini, J. H. (1997). hMre11 and hRad50 nuclear foci are induced during the normal cellular response to DNA double-strand breaks. *Mol.Cell Biol.* **17**, 6087-6096.

Morrison, L. (2004) PhD Thesis. The University of Glasgow.

Masson, J. Y., Davies, A. A., Hajibagheri, N., Van Dyck, E., Benson, F. E., Stasiak, A. Z., Stasiak, A. and West, S. C. (1999). The meiosis-specific recombinase hDmc1 forms ring structures and interacts with hRad51. *EMBO J.* **18**, 6552-6560.

Masson, J. Y., Stasiak, A. Z., Stasiak, A., Benson, F. E. and West, S. C. (2001a). Complex formation by the human RAD51C and XRCC3 recombination repair proteins. *Proc.Natl.Acad.Sci.U.S.A* **98**, 8440-8446.

Masson, J. Y., Tarsounas, M. C., Stasiak, A. Z., Stasiak, A., Shah, R., McIlwraith, M. J., Benson, F. E. and West, S. C. (2001b). Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev.* **15**, 3296-3307.

Masson, J. Y. and West, S. C. (2001). The Rad51 and Dmc1 recombinases: a non-identical twin relationship. *Trends Biochem.Sci.* **26**, 131-136.

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

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

Matthews, K. R., 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.

Mazin, A. V., Zaitseva, E., Sung, P. and Kowalczykowski, S. C. (2000). Tailed duplex DNA is the preferred substrate for Rad51 protein-mediated homologous pairing. *EMBO J.* **19**, 1148-1156.

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 conversions 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.

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

McKean, P. G., Keen, J. K., Smith, D. F. and Benson, F. E. (2001). Identification and characterisation of a RAD51 gene from *Leishmania major*. *Mol.Biochem.Parasitol.* **115**, 209-216.

McKim, K. S. and Hayashi-Hagihara, A. (1998). mei-W68 in *Drosophila melanogaster* encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved. *Genes Dev.* **12**, 2932-2942.

Melo, J. A., Cohen, J. and Toczyski, D. P. (2001). Two checkpoint complexes are independently recruited to sites of DNA damage in vivo. *Genes Dev.* **15**, 2809-2821.

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.

Melville, S. E., Leech, V., Navarro, M. and Cross, G. A. (2000). The molecular karyotype of the megabase chromosomes of *Trypanosoma brucei* stock 427. *Mol.Biochem.Parasitol.* **111**, 261-273.

Mercier, G., Denis, Y., Marc, P., Picard, L. and Dutreix, M. (2001). Transcriptional induction of repair genes during slowing of replication in irradiated *Saccharomyces cerevisiae*. *Mutat.Res.* **487**, 157-172.

Michel, B., Grompone, G., Flores, M. J. and Bidnenko, V. (2004). Multiple pathways process stalled replication forks. *Proc.Natl.Acad.Sci.U.S.A* **101**, 12783-12788.

Mimori, T. and Hardin, J. A. (1986). Mechanism of interaction between Ku protein and DNA. *J.Biol.Chem.* **261**, 10375-10379.

Moore, J. K. and Haber, J. E. (1996). Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. *Mol.Cell Biol.* **16**, 2164-2173.

Namsaraev, E. A. and Berg, P. (1998). Branch migration during Rad51-promoted strand exchange proceeds in either direction. *Proc.Natl.Acad.Sci.U.S.A* **95**, 10477-10481.

Namsaraev, E. A. and Berg, P. (2000). Rad51 uses one mechanism to drive DNA strand exchange in both directions. *J.Biol.Chem.* **275**, 3970-3976.

Nash, T. E. (2002). Surface antigenic variation in *Giardia lamblia*. *Mol.Microbiol.* **45**, 585-590.

Nassif, N., Penney, J., Pal, S., Engels, W. R. and Gloor, G. B. (1994). Efficient copying of nonhomologous sequences from ectopic sites via P- element-induced gap repair. *Mol.Cell Biol.* **14**, 1613-1625.

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

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

Neyton, S., Lespinasse, F., Moens, P. B., Paul, R., Gaudray, P., Paquis-Flucklinger, V. and Santucci-Darmanin, S. (2004). Association between MSH4 (MutS homologue 4) and the DNA strand-exchange RAD51 and DMC1 proteins during mammalian meiosis. *Mol.Hum.Reprod* **10**, 917-924.

Novak, J. E., Ross-Macdonald, P. B. and Roeder, G. S. (2001). The budding yeast Msh4 protein functions in chromosome synapsis and the regulation of crossover distribution. *Genetics* **158**, 1013-1025.

O'Regan, P., Wilson, C., Townsend, S. and Thacker, J. (2001). XRCC2 is a nuclear RAD51-like protein required for damage-dependent RAD51 focus formation without the need for ATP binding. *J.Biol.Chem.* **276**, 22148-22153.

Osakabe, K., Yoshioka, T., Ichikawa, H. and Toki, S. (2002). Molecular cloning and characterization of RAD51-like genes from *Arabidopsis thaliana*. *Plant Mol.Biol.* **50**, 71-81.

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput.Appl.Biosci.* **12**, 357-358.

Page, Scott L. and Hawley, R. Scott (2003). Chromosome Choreography: The Meiotic Ballet. *Science* **301**, 785-789.

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

Passy, S. I., Yu, X., Li, Z., Radding, C. M., Masson, J. Y., West, S. C. and Egelman, E. H. (1999). Human Dmc1 protein binds DNA as an octameric ring. *Proc.Natl.Acad.Sci.U.S.A* **96**, 10684-10688.

Pastwa, E. and Blasiak, J. (2003). Non-homologous DNA end joining. *Acta Biochim.Pol.* **50**, 891-908.

Paull, T. T. and Gellert, M. (1998). The 3' to 5' exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks. *Mol.Cell* **1**, 969-979.

Pays, E. (1989). Pseudogenes, chimaeric genes and the timing of antigen variation in African trypanosomes. *Trends Genet.* **5**, 389-391.

Pays, E., Houard, S., Pays, A., Van Assel, S., Dupont, F., Aerts, D., Huet-Duvillier, G., Gomes, V., Richet, C., Degand, P. and . (1985). *Trypanosoma brucei*: the extent of conversion in antigen genes may be related to the DNA coding specificity. *Cell* **42**, 821-829.

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

Pays, E., Vanhamme, L. and Berberof, M. (1994). Genetic controls for the expression of surface antigens in African trypanosomes. *Annu.Rev.Microbiol.* **48**, 25-52.

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

Pellegrini, L. and Venkitaraman, A. (2004). Emerging functions of BRCA2 in DNA recombination. *Trends Biochem.Sci.* **29**, 310-316.

Pellegrini, L., Yu, D. S., Lo, T., Anand, S., Lee, M., Blundell, T. L. and Venkitaraman, A. R. (2002). Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* **420**, 287-293.

Petukhova, G., Stratton, S. and Sung, P. (1998). Catalysis of homologous DNA pairing by yeast Rad51 and Rad54 proteins. *Nature* **393**, 91-94.

Petukhova, G., Stratton, S. A. and Sung, P. (1999a). Single strand DNA binding and annealing activities in the yeast recombination factor Rad59. *J.Biol.Chem.* **274**, 33839-33842.

Petukhova, G., Van Komen, S., Vergano, S., Klein, H. and Sung, P. (1999b). Yeast Rad54 promotes Rad51-dependent homologous DNA pairing via ATP hydrolysis-driven change in DNA double helix conformation. *J.Biol.Chem.* **274**, 29453-29462.

Raderschall, E., Golub, E. I. and Haaf, T. (1999). Nuclear foci of mammalian recombination proteins are located at single- stranded DNA regions formed after DNA damage. *Proc.Natl.Acad.Sci.U.S.A* **96**, 1921-1926.

Rao, B. J., Chiu, S. K., Bazemore, L. R., Reddy, G. and Radding, C. M. (1995). How specific is the first recognition step of homologous recombination? *Trends Biochem.Sci.* **20**, 109-113.

Rayssiguier, C., Thaler, D. S. and Radman, M. (1989). The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* **342**, 396-401.

Reich, C. I., McNeil, L. K., Brace, J. L., Brucker, J. K. and Olsen, G. J. (2001). Archaeal RecA homologues: different response to DNA-damaging agents in mesophilic and thermophilic Archaea. *Extremophiles.* **5**, 265-275.

Ristic, D., Modesti, M., Kanaar, R. and Wyman, C. (2003). Rad52 and Ku bind to different DNA structures produced early in double-strand break repair. *Nucleic Acids Res.* **31**, 5229-5237.

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.

Robinson, N. P., McCulloch, R., Conway, C., Browitt, A. and Barry, J. D. (2002). Inactivation of Mre11 Does Not Affect VSG Gene Duplication Mediated by Homologous Recombination in *Trypanosoma brucei*. *J.Biol.Chem.* **277**, 26185-26193.

Roditi, I. and Liniger, M. (2002). Dressed for success: the surface coats of insect-borne protozoan parasites. *Trends Microbiol.* **10**, 128-134.

Roeder, G. S. (1997). Meiotic chromosomes: it takes two to tango. *Genes Dev.* **11**, 2600-2621.

Romanienko, P. J. and Camerini-Otero, R. D. (2000). The mouse Spo11 gene is required for meiotic chromosome synapsis. *Mol.Cell* **6**, 975-987.

Rudenko, G. (2000). The polymorphic telomeres of the African Trypanosome *trypanosoma brucei*. *Biochem.Soc.Trans.* **28**, 536-540.

Rudenko, G., Blundell, P. A., Dirks-Mulder, 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 *T. brucei*. *Cell* **83**, 547-553.

Rudenko, G., Blundell, P. A., Taylor, M. C., Kieft, R. and Borst, P. (1994). VSG gene expression site control in insect form *Trypanosoma brucei*. *EMBO J.* **13**, 5470-5482.

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

Salmon, D., Geuskens, M., Hanocq, F., Hanocq-Quertier, J., Nolan, D., Ruben, L. and Pays, E. (1994). A novel heterodimeric transferrin receptor encoded by a pair of VSG expression site-associated genes in *T. brucei*. *Cell* **78**, 75-86.

Schofield, M. J. and Hsieh, P. (2003). DNA mismatch repair: molecular mechanisms and biological function. *Annu.Rev.Microbiol.* **57**, 579-608.

Scholler, J. K., Myler, P. J. and Stuart, K. D. (1989). A novel telomeric gene conversion in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **35**, 11-19.

Schwacha, A. and Kleckner, N. (1997). Interhomolog bias during meiotic recombination: meiotic functions promote a highly differentiated interhomolog-only pathway. *Cell* **90**, 1123-1135.

Sedgwick, B. (2004). Repairing DNA-methylation damage. *Nat.Rev.Mol.Cell Biol.* **5**, 148-157.

Shao, X. and Grishin, N. V. (2000). Common fold in helix-hairpin-helix proteins. *Nucleic Acids Res.* **28**, 2643-2650.

Sheader, K., Berberof, M., Isobe, T., Borst, P. and Rudenko, G. (2003). Delineation of the regulated Variant Surface Glycoprotein gene expression site domain of *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **128**, 147-156.

Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K. and Ogawa, T. (1993). Cloning of human, mouse and fission yeast recombination genes homologous to RAD51 and recA. *Nat.Genet.* **4**, 239-243.

Shinohara, A., Ogawa, H. and Ogawa, T. (1992). Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. *Cell* **69**, 457-470.

Shinohara, M., Gasior, S. L., Bishop, D. K. and Shinohara, A. (2000). Tid1/Rdh54 promotes colocalization of rad51 and dmc1 during meiotic recombination. *Proc.Natl.Acad.Sci.U.S.A* **97**, 10814-10819.

Shivji, M. K. and Venkitaraman, A. R. (2004). DNA recombination, chromosomal stability and carcinogenesis: insights into the role of BRCA2. *DNA Repair (Amst)* **3**, 835-843.

Sibanda, B. L., Critchlow, S. E., Begun, J., Pei, X. Y., Jackson, S. P., Blundell, T. L. and Pellegrini, L. (2001). Crystal structure of an Xrcc4-DNA ligase IV complex. *Nat.Struct.Biol.* **8**, 1015-1019.

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.

Sigurdsson, S., Van Komen, S., Bussen, W., Schild, D., Albala, J. S. and Sung, P. (2001). Mediator function of the human Rad51B-Rad51C complex in Rad51/RPA-catalyzed DNA strand exchange. *Genes Dev.* **15**, 3308-3318.

Sigurdsson, S., Van Komen, S., Petukhova, G. and Sung, P. (2002). Homologous DNA pairing by human recombination factors Rad51 and Rad54. *J.Biol.Chem.* **277**, 42790-42794.

Smith, B. T. and Walker, G. C. (1998). Mutagenesis and more: umuDC and the *Escherichia coli* SOS response. *Genetics* **148**, 1599-1610.

Smith, G. C. and Jackson, S. P. (1999). The DNA-dependent protein kinase. *Genes Dev.* **13**, 916-934.

Smith, G. R., Boddy, M. N., Shanahan, P. and Russell, P. (2003). Fission yeast Mus81.Eme1 Holliday junction resolvase is required for meiotic crossing over but not for gene conversion. *Genetics* **165**, 2289-2293.

Smith, K. N. and Nicolas, A. (1998). Recombination at work for meiosis. *Curr.Opin.Genet.Dev.* **8**, 200-211.

Snowden, T., Acharya, S., Butz, C., Berardini, M. and Fishel, R. (2004). hMSH4-hMSH5 recognizes Holliday Junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes. *Mol.Cell* **15**, 437-451.

Staeva-Vieira, E., Yoo, S. and Lehmann, R. (2003). An essential role of DmRad51/SpnA in DNA repair and meiotic checkpoint control. *EMBO J.* **22**, 5863-5874.

Stevens, J. R. and Gibson, W. (1999). The molecular evolution of trypanosomes. *Parasitol.Today* **15**, 432-437.

Steverding, D., Stierhof, Y. D., Chaudhri, M., Ligtenberg, M., Schell, D., Beck-Sickinger, A. G. and Overath, P. (1994). ESAG 6 and 7 products of *Trypanosoma brucei* form a transferrin binding protein complex. *Eur.J.Cell Biol.* **64**, 78-87.

Story, R. M., Bishop, D. K., Kleckner, N. and Steitz, T. A. (1993). Structural relationship of bacterial RecA proteins to recombination proteins from bacteriophage T4 and yeast. *Science* **259**, 1892-1896.

Story, R. M., Weber, I. T. and Steitz, T. A. (1992). The structure of the E. coli recA protein monomer and polymer. *Nature* **355**, 318-325.

Sugiyama, T., Zaitseva, E. M. and Kowalczykowski, S. C. (1997). A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the *Saccharomyces cerevisiae* Rad51 protein. *J.Biol.Chem.* **272**, 7940-7945.

Sung, P. (1997a). Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. *J.Biol.Chem.* **272**, 28194-28197.

Sung, P. (1997b). Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein A to promote DNA strand exchange by Rad51 recombinase. *Genes Dev.* **11**, 1111-1121.

Sung, P. and 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.

Symington, L. S. (2002). Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol.Mol.Biol.Rev.* **66**, 630-70, table.

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.

Takata, M., Sasaki, M. S., Tachiiri, S., Fukushima, T., Sonoda, E., Schild, D., Thompson, L. H. and Takeda, S. (2001). Chromosome instability and defective recombinational repair in knockout mutants of the five Rad51 paralogs. *Mol.Cell Biol.* **21**, 2858-2866.

Tan, K. S., Leal, S. T. and Cross, G. A. (2002). *Trypanosoma brucei* MRE11 is non-essential but influences growth, homologous recombination and DNA double-strand break repair. *Mol.Biochem.Parasitol.* **125**, 11-21.

Tanaka, K., Kagawa, W., Kinebuchi, T., Kurumizaka, H. and Miyagawa, K. (2002). Human Rad54B is a double-stranded DNA-dependent ATPase and has biochemical properties different from its structural homolog in yeast, Tid1/Rdh54. *Nucleic Acids Res.* **30**, 1346-1353.

Tarsounas, M., Davies, A. A. and West, S. C. (2004a). RAD51 localization and activation following DNA damage. *Philos.Trans.R.Soc.Lond B Biol.Sci.* **359**, 87-93.

Tarsounas, M., Davies, D. and West, S. C. (2003). BRCA2-dependent and independent formation of RAD51 nuclear foci. *Oncogene* **22**, 1115-1123.

Tarsounas, M., Munoz, P., Claas, A., Smiraldo, P. G., Pittman, D. L., Blasco, M. A. and West, S. C. (2004b). Telomere maintenance requires the RAD51D recombination/repair protein. *Cell* **117**, 337-347.

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

Thacker, J. (1999). A surfeit of RAD51-like genes? *Trends Genet.* **15**, 166-168.

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

Timmers, H. T., 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.

Trujillo, K. M., Roh, D. H., Chen, L., Van Komen, S., Tomkinson, A. and Sung, P. (2003). Yeast xrs2 binds DNA and helps target rad50 and mre11 to DNA ends. *J.Biol.Chem.* **278**, 48957-48964.

Trujillo, K. M. and Sung, P. (2001). DNA structure-specific nuclease activities in the Saccharomyces cerevisiae Rad50*Mre11 complex. *J.Biol.Chem.* **276**, 35458-35464.

Turner, C. M., Aslam, N. and Dye, C. (1995). Replication, differentiation, growth and the virulence of *Trypanosoma brucei* infections. *Parasitology* **111 (Pt 3)**, 289-300.

Turner, C. M. and Barry, J. D. (1989). High frequency of antigenic variation in *Trypanosoma brucei rhodesiense* infections. *Parasitology* **99 Pt 1**, 67-75.

Turner, C. M., 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 (Pt 2)**, 269-276.

Tyler, K. M. and Engman, D. M. (2000). Flagellar elongation induced by glucose limitation is preadaptive for *Trypanosoma cruzi* differentiation. *Cell Motil.Cytoskeleton* **46**, 269-278.

Tyler, K. M., Higgs, P. G., Matthews, K. R. and Gull, K. (2001). Limitation of *Trypanosoma brucei* parasitaemia results from density- dependent parasite differentiation and parasite killing by the host immune response. *Proc.R.Soc.Lond B Biol.Sci.* **268**, 2235-2243.

Ulbert, S., Chaves, I. and Borst, P. (2002a). Expression site activation in *Trypanosoma brucei* with three marked variant surface glycoprotein gene expression sites. *Mol.Biochem.Parasitol.* **120**, 225-235.

Ulbert, S., Cross, M., Boorstein, R. J., Teebor, G. W. and Borst, P. (2002b). Expression of the human DNA glycosylase hSMUG1 in *Trypanosoma brucei* causes DNA damage and interferes with J biosynthesis. *Nucleic Acids Res.* **30**, 3919-3926.

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

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

van der Woude, M. W. and Baumler, A. J. (2004). Phase and antigenic variation in bacteria. *Clin.Microbiol.Rev.* **17**, 581-611, table.

Van Dyck, E., Stasiak, A. Z., Stasiak, A. and West, S. C. (1999). Binding of double-strand breaks in DNA by human Rad52 protein. *Nature* **398**, 728-731.

van Gent, D. C., Hoeijmakers, J. H. and Kanaar, R. (2001). Chromosomal stability and the DNA double-stranded break connection. *Nat.Rev.Genet.* **2**, 196-206.

Van Komen, S., Petukhova, G., Sigurdsson, S., Stratton, S. and Sung, P. (2000). Superhelicity-driven homologous DNA pairing by yeast recombination factors Rad51 and Rad54. *Mol.Cell* **6**, 563-572.

Van Komen, S., Petukhova, G., Sigurdsson, S. and Sung, P. (2002). Functional cross-talk among Rad51, Rad54, and replication protein A in heteroduplex DNA joint formation. *J.Biol.Chem.* **277**, 43578-43587.

van Leeuwen, F., Kieft, R., Cross, M. and Borst, P. (2000). Tandemly repeated DNA is a target for the partial replacement of thymine by beta-D-glucosyl-hydroxymethyluracil in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **109**, 133-145.

van Leeuwen, F., Taylor, M. C., Mondragon, A., Moreau, H., Gibson, W., Kieft, R. and Borst, P. (1998). beta-D-glucosyl-hydroxymethyluracil is a conserved DNA modification in kinetoplastid protozoans and is abundant in their telomeres. *Proc.Natl.Acad.Sci.U.S.A* **95**, 2366-2371.

van Leeuwen, F., Wijsman, E. R., Kieft, R., van der Marel, G. A., van Boom, J. H. and Borst, P. (1997). Localization of the modified base J in telomeric VSG gene expression sites of *Trypanosoma brucei*. *Genes Dev.* **11**, 3232-3241.

van Leeuwen, F., Wijsman, E. R., Kuyl-Yeheskiely, E., van der Marel, G. A., van Boom, J. H. and Borst, P. (1996). The telomeric GGGTTA repeats of *Trypanosoma brucei* contain the hypermodified base J in both strands. *Nucleic Acids Res.* **24**, 2476-2482.

Vanhamme, L., Berberof, M., Le Ray, D. and Pays, E. (1995). Stimuli of differentiation regulate RNA elongation in the transcription units for the major stage-specific antigens of *Trypanosoma brucei*. *Nucleic Acids Res.* **23**, 1862-1869.

Vanhamme, L., Lecordier, L. and Pays, E. (2001a). Control and function of the bloodstream variant surface glycoprotein expression sites in *Trypanosoma brucei*. *Int.J.Parasitol.* **31**, 523-531.

Vanhamme, L., Paturiaux-Hanocq, F., Poelvoorde, P., Nolan, D. P., Lins, L., Van den, Abbeele J., Pays, A., Tebabi, P., Van Xong, H., Jacquet, A., Moguilevsky, N., Dieu, M., Kane, J. P., De Baetselier, P., Brasseur, R. and Pays, E. (2003). Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature* **422**, 83-87.

Vanhamme, L. and Pays, E. (1995). Control of gene expression in trypanosomes. *Microbiol.Rev.* **59**, 223-240.

Vanhamme, L., Pays, E., McCulloch, R. and Barry, J. D. (2001b). An update on antigenic variation in African trypanosomes. *Trends Parasitol.* **17**, 338-343.

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

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

Vassella, E., Reuner, B., Yutzy, B. and Boshart, M. (1997). Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. *J.Cell Sci.* **110 (21)**, 2661-2671.

Vengrova, S. and Dalgaard, J. Z. (2004). RNase-sensitive DNA modification(s) initiates *S. pombe* mating-type switching. *Genes Dev.* **18**, 794-804.

Walker, G. C. (1984). Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol.Rev.* **48**, 60-93.

Walker, J. E., Saraste, M., Runswick, M. J. and Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945-951.

Wang, T. F., Kleckner, N. and Hunter, N. (1999). Functional specificity of MutL homologs in yeast: evidence for three Mlh1-based heterocomplexes with distinct roles during meiosis in recombination and mismatch correction. *Proc.Natl.Acad.Sci.U.S.A* **96**, 13914-13919.

Weller, G. R., Kysela, B., Roy, R., Tonkin, L. M., Scanlan, E., Della, M., Devine, S. K., Day, J. P., Wilkinson, A., d'Adda, di Fagagna, Devine, K. M., Bowater, R. P., Jeggo, P. A., Jackson, S. P. and Doherty, A. J. (2002). Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* **297**, 1686-1689.

West, R. B., Yaneva, M. and Lieber, M. R. (1998). Productive and nonproductive complexes of Ku and DNA-dependent protein kinase at DNA termini. *Mol.Cell Biol.* **18**, 5908-5920.

West, S. C. (1996). The RuvABC proteins and Holliday junction processing in *Escherichia coli*. *J.Bacteriol.* **178**, 1237-1241.

West, S. C. (1997). Processing of recombination intermediates by the RuvABC proteins. *Annu.Rev.Genet.* **31**, 213-244.

Weterings, E. and van Gent, D. C. (2004). The mechanism of non-homologous end-joining: a synopsis of synapsis. *DNA Repair (Amst)* **3**, 1425-1435.

Wiese, C., Collins, D. W., Albala, J. S., Thompson, L. H., Kronenberg, A. and Schild, D. (2002). Interactions involving the Rad51 paralogs Rad51C and XRCC3 in human cells. *Nucleic Acids Res.* **30**, 1001-1008.

Wilson, T. E., Topper, L. M. and Palmbos, P. L. (2003). Non-homologous end-joining: bacteria join the chromosome breakdance. *Trends Biochem.Sci.* **28**, 62-66.

Wirtz, E., Leal, S., Ochatt, C. and Cross, G. A. (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. *Mol.Biochem.Parasitol.* **99**, 89-101.

Xong, H. V., Vanhamme, L., Chamekh, M., Chimfwembe, C. E., Van den, Abbeele J., Pays, A., Van Meirvenne, N., Hamers, R., De Baetselier, P. and Pays, E. (1998). A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense*. *Cell* **95**, 839-846.

Yamaguchi-Iwai, Y., Sonoda, E., Buerstedde, J. M., Bezzubova, O., Morrison, C., Takata, M., Shinohara, A. and Takeda, S. (1998). Homologous recombination, but not DNA repair, is reduced in vertebrate cells deficient in RAD52. *Mol.Cell Biol.* **18**, 6430-6435.

Yamaguchi-Iwai, Y., Sonoda, E., Sasaki, M. S., Morrison, C., Haraguchi, T., Hiraoka, Y., Yamashita, Y. M., Yagi, T., Takata, M., Price, C., Kakazu, N. and Takeda, S. (1999). Mre11 is essential for the maintenance of chromosomal DNA in vertebrate cells. *EMBO J.* **18**, 6619-6629.

Yang, S., VanLoock, M. S., Yu, X. and Egelman, E. H. (2001). Comparison of bacteriophage T4 UvsX and human Rad51 filaments suggests that RecA-like polymers may have evolved independently. *J.Mol.Biol.* **312**, 999-1009.

Yokoyama, H., Kurumizaka, H., Ikawa, S., Yokoyama, S. and Shibata, T. (2003). Holliday junction binding activity of the human Rad51B protein. *J.Biol.Chem.* **278**, 2767-2772.

Yoshida, K., Kondoh, G., Matsuda, Y., Habu, T., Nishimune, Y. and Morita, T. (1998). The mouse RecA-like gene Dmc1 is required for homologous chromosome synapsis during meiosis. *Mol.Cell* **1**, 707-718.